Essays on Language, Cognition, and Group Decision-Making



Laura Fritsch

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2023

Contents

1. Acknowledgements **3**

2. Introduction **5**

3. Meaning as Use: Linguistic Foundations for the Behavioral Theory of the Firm **9**

3.1. Introduction 10

3.2. The Behavioral Theory of the Firm 12

3.3. Language Games: Meaning as Use and Family Resemblances 18

3.4. A Model of Organizational Categorization and Action 20

3.5. Discontinuous Organizational Change 32

3.6. Case Studies 36

3.7. Implications for the Behavioral Theory of the Firm 42

3.8. Proposed Empirical Tests 46

3.9. Conclusion 50

4. Vocabularies, Attention, and Imprinting in Organizations: An Experimental Approach **52**

4.1. Introduction 53

4.2. Literature Review 55

4.3. Experiment 67

4.4. Results 71

4.5. Conclusion 100

4.6. Implications 101

4.A. Appendix: Tables 104

4.B. Appendix: Experiment Details 108

5. Market Reactions to Gendered Speech Patterns: Uptalk, Earnings Calls, and the #MeToo Movement **114**

5.1. Introduction . . . 114

5.2. Literature Review . . . 120

5.3. Data and Methodology . . . 126

5.4. Uptalk . . . 129

5.5. Market Reactions . . . 133

5.6. The #MeToo Movement . . . 147

5.7. Conclusion . . . 156

5.A. Appendix . . . 158

6. Analysts Downgrade Stock Recommendations When Female CEOs Use Uptalk **175**

6.1. Introduction . . . 176

6.2. Literature Review . . . 178

6.3. Data Description . . . 181

6.4. Research Design . . . 183

6.5. Results . . . 184

6.6. Discussion and Implications . . . 192

6.A. Appendix . . . 195

7. Conclusion **200**

7.1. Integrative Reflections . . . 200

7.2. Future Research . . . 203

1

Acknowledgements

First and foremost, I would like to extend my deepest gratitude to my supervisor, Professor Alan Morrison, for encouraging all my insane research ideas while keeping me sane during a global pandemic. Throughout the past five years, I have counted myself lucky to be the recipient of his unwavering support, kindness, and guidance. I am even luckier, however, to have found a cherished friend. In short, I would not be the researcher and person I am today without Alan. "O captain! My captain".

I would also like to thank my co-authors: Professor Anantha Divakaruni, Professor Aharon Cohen Mohliver, Professor Brian Gunia, Professor Howard Jones, and Professor Eero Vaara. They taught me that in academia, as in life, we should always stand on the shoulders of giants.

My heartfelt thanks to my doctoral committee members, Professor Sameer Srivastava and Professor Marya Besharov, for reading my excruciatingly long dissertation and giving me incredibly useful guidance.

I also owe a great debt of gratitude to all my academic mentors and colleagues: Professor Paul Gouvard, Professor Michael Jensen, Professor William Ocasio, Professor Ken Okamura, Professor Ann Tenbrunsel, Professor Linda Treviño, Professor Marc Ventresca, and Professor Mungo Wilson. Your invaluable comments and guidance have profoundly shaped my academic journey.

A special thanks to Rupert Younger, the Director of the Centre for Corporate Reputation, for introducing me to my supervisor, funding my doctorate, and championing my academic work. It would not be an overstatement to say that you changed the course of my career. I would also like to thank Hannah Morgan for being the backbone of the Centre for Corporate Reputation and making

sure that we get to and from conferences in one piece.

I will be forever thankful for the CCR Doctoral Scholarship, the Alan Turing Institute Doctoral Scholarship, the European and Social Research Council Scholarship, and the Scatcherd European Scholarship. Without these generous awards, this doctorate would not have been possible.

To my doctoral cohort—Dr. Seth Jenson, Dr. Javier Amaya, David Lehman, Daniel Pesch, and Wayne Lim—I am immensely thankful for your camaraderie over the past five years and cherish the bonds we have formed. Great company in a journey makes the way seem shorter.

I would like to wholeheartedly thank my friends and siblings for keeping me sane: Gabriel Fritsch, Felipe Fritsch, Silvia Correia, Dr. Katie Hoeger, Dr. Adriano Fernandes, Dr. Nicolas Lippolis, Dr. Julian Saadi-Lalande, Dr. Jishnu Mukhoti, Dr. Caitlyn McGeer, Claudia Whyte, Laurelle Remzi, Noyan Songur, Irene Pontisso, Kayvane Shakerifar, Antonia Klabin, Trisha Paul, Amanda Muniz, Larissa Guimarães, and Clara Motta.

To Lupin, my loyal canine companion, thank you for being a good boy and providing me with much-needed perspective.

A colossal thank you to my parents who taught me the value of education and worked tirelessly to give me every opportunity. To my father, thank you for planting the idea that I could move halfway across the world in pursuit of my dreams. I hope you have forgiven me for choosing Oxford instead of the other place. To my mother, thank you for being my North Star. She gave up her own PhD to raise me so, in every way, this one is for both of us.

At last, and forever truly first, I would like to thank my fiancé, Dr. Roman Rothaermel, who has taught me that to be loved is to be seen. Throughout this doctoral odyssey, he has kept me fed and watered, anticipated needs I didn't even know I had, believed in me when I didn't believe in myself, picked me up when I was down, taught me how to write my first line of code, and closed my computer when I lacked the self-discipline. Ich liebe dich.

2

Introduction

It is impossible to imagine social life without language. People communicate in order to get things done; the way that they communicate therefore reflects their understanding of how the world works, and it also conveys to other people how they should understand the world. The meanings that language conveys are created within groups of people and, hence, studying the way that language is used can shed light upon fundamental questions about group cognition, group identity, and other important cultural phenomena.

This thesis examines the micro-level use of language in organizations, and so sheds new light upon the way that groups create and manipulate meaning, and upon what it means for language and meaning to be explicitly organizational, rather than personal. The thesis starts from Wittgenstein's (1953) insight that meaning is use: that is, that language is a means of getting things done and, hence, that meaning is something that exists in the stream of life and not as a private mental model. By this account, language can be understood in organizations and other groups only with reference to the uses to which it is put. The thesis uses this insight to examine the way in which meaning is constructed as organizations get things done; to build a micro-foundation for aspects of the behavioural theory of the firm and to understand how gender-typed speech can have real-economic effects and impact analysts' assessments of newly appointed female CEOs.

The thesis thus makes a conceptual contribution to the study of language and culture in organizations. The thesis presents a new theory that explains how language, categorization, and organizational action interact with one another to create persistent organizational behaviours. The same theory can explain how organizations search and identify the set of possible responses to

new technological or social cues. It also includes a theoretical framework that sheds the light on the ways in which language becomes imprinted on organizations and, hence, how founders have long-lived effects on the behaviour of their firms beyond their tenure.

Moreover, the thesis makes several methodological innovations. First, the thesis presents an experimental study that is novel in management and economics literatures of the persistence and evolution of language. Second, the thesis presents one of the first analyses that uses audio recordings to study the effect of gender in social situations. More precisely, it presents the first study of uptalk in earnings calls and its effect upon market prices. These methods enable the analysis of very large datasets: the experiment in Chapter 4 generated a high-volume of text that is usual in laboratory settings; and the earnings call dataset used in Chapter 5 and Chapter 6 comprises around 50 million minutes of calls.

Chapter 3 presents a linguistic microfoundation for the Behavioural Theory of the Firm (Cyert and March, 1963). It uses Wittgenstein's (1953) categorization of "meaning as use," first, to argue that organizations build categorical models of their environments; and, second, that they attach use-meanings to those categories. The chapter presents a process model in which those activities can become decoupled and, hence, organizational routines naturally deviate from classical means-end rationality. The chapter conceptualizes *language work*, which is the process by which organizations consciously deliberate over categories and their meanings. The chapter embeds categorization and language work in a process model from which many important aspects of the Behavioural Theory of the Firm naturally emerge. In addition, the chapter presents a natural characterization of technological adjacency that is organization-specific and also historically contingent. The model is used to explain two case studies involving social and technological change. The chapter concludes by proposing some testable hypotheses based on the model.

Chapter 4 studies how vocabularies can microfound organizational imprinting. It presents a process model within which organizations converge on specific vocabularies that enable them to perform tasks efficiently and, in turn, those vocabularies support commonly understood categories that enable organizational members to make sense of the world. Once imprinted, they are long-lived and affect the ways in which organizational members perceive their world and, hence, upon

their actions. The chapter studies these phenomena in an experimental setting within which participants communicate orally to perform an image recognition task. Natural Language Processing (NLP) is performed on transcripts of the subjects' speech in order to quantify the formation, transmission, and persistence of organizational vocabularies. When subjects leave the experiment, they complete a basic memory task that measures the way that their attention was directed during the experiment. This set-up therefore facilitates an analysis of the relationship between strategic vocabularies and organizational imprinting. The chapter yields five main results. First, the repeated iteration of tasks by organizational actors imprints the vocabulary on the organization. Second, once a vocabulary has been imprinted, organizations operate more efficiently. Third, holding personnel, processes, categories, and goals fixed, organizational identity affects the vocabularies that people use to perform tasks. Fourth, the probability that a word in a vocabulary affects a organization member's cognition of their environment is affected by the way that the member encounters the vocabulary. Finally, the vocabularies established by founders have persistent cognitive impacts on organizational members.

Chapter 5 analyses the way that gender-typed speech by people with private information affects price formation. The chapter introduces a novel concept into the economics and finance literatures: uptalk. Uptalk occurs when the end of a speech ends has a higher acoustic tone than its body. Uptalk is more likely to occur when women speak than when men do, and that it is associated with a lack of speaker confidence. The chapter uses audio recordings of earnings calls of public US firms to measure uptalk, and relate it to post-call analyst recommendations and to realised earnings. This yields four main results. First, markets respond negatively when female executives exhibit unexpectedly high levels of uptalk: that is, when women exhibit high levels of a stereotypically female speech pattern. This result manifests in a higher likelihood of downgraded analyst recommendations after a call in which female executives use unexpectedly high levels of uptalk; the same effect does not obtain when male executives exhibit this speech pattern. Second, unexpected uptalk by female executives is associated with lower future earnings, while there is no such statistical relationship between unexpected male executive uptalk and earnings. Strikingly, the data suggests that markets actually under-respond to female executive uptalk. Third, the chapter uses real-time

bid-ask spread data to show that market risk perceptions increase when female executives speak during earnings calls. The chapter suggests that this might be because, if the executive exhibits uptalk, then she reveals potentially bad news about her firm's earnings prospects. Finally, it shows that the #MeToo social movement had a real effect on markets. #MeToo had no effect on market perceptions of female executives but resulted in more positive market responses to male executive uptalk.

Finally, Chapter 6 focuses specifically on CEO transitions and studies the impact of gender and gender-typed speech characteristics on equity analysts' stock recommendations for their respective companies. Examining 753 CEO transitions in US public firms from 2011 to 2019, the chapter finds that, on average, equity analysts downgrade their recommendations following the initial earnings calls led by a newly appointed female CEO who employs high levels uptalk during these calls. No such downgrades are observed when incoming female CEOs exhibit lower levels of uptalk or when incoming male CEOs use uptalk to any degree. Taken together, the chapter concludes that the persistent downgrade in analyst recommendations for incoming female CEOs may be driven by a subtle interaction between gender and gender-typed speech, an explanation consistent with gender role norms. These findings demonstrate the importance of analysing voice data to gain deeper insights into the persistent "gender penalty" experienced by women in financial markets.

3

Meaning as Use: Linguistic Foundations for the Behavioral Theory of the Firm

(Joint with Alan D. Morrison[1])

Abstract

The Behavioral Theory of the Firm has generated deep insights into the ways that individual actions shape, and are shaped by, organizational structures. However, almost sixty years after the theory was introduced by Cyert and March (1963), there is still no agreed-upon microfoundation that bridges the gap between the internal workings of the firm and its engagement with the external environment. This article suggests that, in order to find that bridge, researchers should take a linguistic turn. The argument starts from Wittgenstein's (1953) insight that linguistic meaning derives from social use rather than from private mental models or from more foundational algorithms. We deploy this insight to explain, first, how organizations build categorical models of their environments, and, second, how they attach use-meanings to those categories. We present a process model in which those activities can become decoupled and, hence, organizational routines naturally deviate from classical means-end rationality. Organizations may need to deliberate over categories and their meanings. We model the associated *language work*. In our framework, firms search their environments when they feel to meet their aspirations; their searches start locally and cease when they are able to meet their aspirations; such searches are "problemistic." Our model yields a natural characterization of techno-

[1]The authors thank Michael Jensen and Marc Ventresca for their comments on an earlier draft, as well as audiences at the the Academy of Management Conference 2022 and 2023; the European Group for Organizational Studies Conference 2022; and at brown-bag seminars at Oxford.

logical adjacency that is organization-specific and also historically contingent. We show how our model explains two case studies involving social and technological change.

3.1. Introduction

The Behavioral Theory of the Firm starts from the fact that organizations seldom act optimally in light of their goals and the relevant facts (Cyert and March, 1963; March and Simon, 1958; March and Simon, 1958). Indeed, organizations frequently structure the social interactions of their members so as to render optimal decision making difficult, if not impossible. For example, organizations frame their members' perceptions of the external environment (Gavetti, 2005; Gavetti and Levinthal, 2000; Levinthal and March, 1993), direct those members' attention (Ocasio, 2011; Ocasio and Joseph, 2005; Ocasio, 1997; March and Olsen, 1976), and rely upon work routines that restrict individual freedom of choice (Levinthal and Rerup, 2006; Papadakis, Lioukas, and Chambers, 1998). These facts affect the way in which organizations search for problems to practical solutions (Rivkin and Siggelkow, 2003; Rivkin, 2000; Levinthal, 1997), and the satisficing criteria used to decide when a search has ended (Nelson and Winter, 1982; Cyert and March, 1963; March and Simon, 1958).

Notwithstanding the enormous importance of the Behavioural Theory of the Firm, it continues to lack a micro foundation. Individuals in organizations do not maximize, and it is widely accepted that they do not use a rational instrumental calculus to simplify their environment in light of the costs of becoming informed and processing information. The literature identifies some important facts about the ways in which they *do* behave. First, actors satisfice with respect to some institutional standard (Nelson and Winter, 1982; Cyert and March, 1963; March and Simon, 1958). Second, organizations search their environments only when prompted to do so by a perceived problem, and they stop searching when they identify a solution that fits an organizational pattern (Cyert and March, 1963). Third, people inside organizations naturally follow rules and standard procedures that shape behaviour in path-dependent, but not necessarily rational, ways. Nonetheless, it has proved extremely hard to identify a framework that can micro found all of these facts.

This article presents an account of organizational behaviour that depends upon the construction

and management of organizational languages. The key to our approach is the insight that the meaning of a language exists *between* its speakers and not in mental images that the language attempts to convey. That is, in our model, "meaning is use" (Wittgenstein, 1953). This approach allows us to generate a linguistic explanation for organizational behaviour. Our model shows how language mediates organizational adaptation and organizational learning. We show how language can underpin the most persistent organizational routines, and how those routines can cause the organization to become decoupled from means-end rationality. We show that organizational responses to social and technological changes are essentially linguistic. We show how language work can result in dramatic, discontinuous, and unexpected changes to organizational behaviour. The linguistic approach thus generates predictions regarding organizational learning, attention direction, organizational adaptation, and the evolution of organizational routines.

Our analysis uses a process model that links environmental cues to organizational actions. In the model, cues are assigned to organizational categories. This move extends a prior literature in which categorization was associated with the cognition of individual people. The assignment process relies upon family resemblances that are perceived by the relevant organizational communities and, hence, is linguistic. The meaning of those categories is the use to which they are put. Our model therefore assigns meaning in two stages: first, by identifying family resemblances and, second, by taking actions. The two stages become routine; they are separated in the organization and can become decoupled in ways that violate standard economic precepts of rationality. Changes to either stage of meaning creation require conscious deliberation and action that we term *language work*. We show how language work depends upon organizational goals in path-dependent ways. In our model, language work determines which technologies an organization adopts and which it ignores. Technologies are adopted when they are *adjacent* to the organization in a sense that we make precise. Our work therefore yields a concrete conceptualization of the "adjacent possibles" of Stuart Kauffman (2002).

We use the predictions generated by our analysis to shed light on two case studies of discontinuous organizational change. The first is social: we argue that the the massive upheaval caused by the #MeToo movement can be explained as an unforeseen consequence of relatively small delib-

erate changes to organizational categorization. The second is technological: we revisit the effects of digital image capture on the photographic industry, and claim that the relative performance of Kodak and Fujifilm in response to that change can be explained in categorical terms.

3.2. The Behavioral Theory of the Firm

The Behavioral Theory of the Firm has generated deep insights into the boundedly rational choices that organizations make and their effect upon organizational outcomes (Cyert and March, 1963). In order to motivate our analysis, this section outlines the basic tenets of the theory. We do not attempt a comprehensive review of the associated literature, which is vast.

The Behavioral Theory of the Firm studies the ways in which organizations and their members affect one another. According to the theory, organizational goals and actions arise from conflict over organizational resources and decision-making power between coalitions of members with differing preferences, identities, and access to information. In turn, organizations shape the preferences and the information-processing capabilities of their members. The central building blocks of the theory are attention directedness, organizational goals and aspirations, problemistic search, routines, and learning; we outline the theory as it relates to each of these ideas,

3.2.1. Attention directedness

Organizations get things done by directing their members' attention on salient facts about the world (Ocasio, 2011; Augier and Prietula, 2007; Ocasio and Joseph, 2005). Attention direction could arise for historic reasons and, hence, need not be deliberate (Dearborn and Simon, 1958). When actors are induced to ignore aspects of their environment, they may deploy their limited processing power more efficiently (March and Olsen, 1976); but they may equally overlook important changes to the world (Gavetti, Greve, Levinthal, and Ocasio, 2012; Levinthal and March, 1993).

Early work on the Behavioral Theory of the Firm argued that attention is directed by the organization's recognized goals (March, 1988; Cyert and March, 1963). More recently, Ocasio's Ocasio (1997) attention-based view of the firm argues that attention directedness should by much

more broadly understood: it could also be shaped by formal and informal structures (Ocasio and Joseph, 2008; Ketokivi and Castaner, 2004), identity (Hoffman and Ocasio, 2001), cognitive models (Kaplan, 2008; Hansen and Haas, 2001), and logics of action (Thornton and Ocasio, 1999; Jacobides, 2007). Each of these approaches can be conceived as "top-down" approach to attention direction; "bottom-up" approaches are rarer (but see Hansen and Haas (2001)), perhaps because such an approach would require a micro-foundation for the Behavioral Theory of the Firm.

3.2.2. Goals and aspirations

In contrast to the models of neoclassical economics, the Behavioral Theory of the Firm claims that organizational goals are the consequence of a continuous process of political contestation (Simon, 1964; Cyert and March, 1963; March, 1962). Firms assess their success or failure in pursuit of their goals with respect to aspirational levels of performance that are determined by looking backwards at the past performance of the firm and its competitors (Tyler and Caner, 2016; Lim and Mccann, 2013; Boyle and Shapira, 2012; Washburn and Bromiley, 2012; Iyer and Miller, 2008; Audia and Greve, 2006; Miller and Chen, 2004; Greve, 1995; Lant, 1992). According to the Behavioral Theory of the Firm, goals and aspirations are updated in response to experience, and not an updated assessment of future states of the world (Lant and Shapira, 2008). In this context, the frames used to interpret experience matter: whether an outcome is good or bad depends critically upon the organization's perceived goals (Levinthal and Rerup, 2021; March, 1994). When an experience is interpreted as a failure, it triggers search processes and organizational change.

3.2.3. Problemistic search

Problemistic search is one of the cornerstones of the Behavioral Theory of the Firm. The theory holds that organizations do not continuously optimise. Instead, they seek solutions to problems that become salient as a result of failures to meet performance aspirations (Levinthal and Rerup, 2021; Posen, Keil, Kim, and Meissner, 2018; Gavetti, Greve, Levinthal, and Ocasio, 2012; Argote and Greve, 2007; Cyert and March, 1963). Moreover, firms do not systematically consider

every possible solution when a search is triggered; instead, they start with "local" possible solutions that are close to their current practices (Levinthal and Rerup, 2021; Simon, 1964), and they stop searching as soon as they find a solution that satisfices (Posen and Chen, 2013; Greve, 2003; Audia, Locke, and Smith, 2000; Simon, 1955). These statements are supported by a substantial empirical literature (Gavetti, Greve, Levinthal, and Ocasio, 2012; Shinkle, 2012; Haleblian, Kim, and Rajagopalan, 2006; Audia, Locke, and Smith, 2000; Greve, 1998; Miller and Chen, 1994; Lant, 1992).

Problemistic search underpins an account of organizations that are inevitably boundedly rational: they do not continuously optimize; they change only when they face an immediate performance problem; and they do not consider the whole of the possible solution space. Organizations in this account are sometimes viewed as automatons, and, by some accounts, the role of deliberation is undertheorized (Posen, Keil, Kim, and Meissner, 2018). Similarly, some authors have called for theories within which R&D efforts and other changes of organizational scope need not be triggered by underperformance (Posen, Keil, Kim, and Meissner, 2018; Greve, 2003; Levinthal, 1997; Levinthal and March, 1981).

3.2.4. Routines

Routines are semi-automatic action blueprints. They are important in the Behavioral Theory of the Firm as mezzo-level phenomena that structure the interaction between organizational members and between those members and the world (Argote and Greve, 2007; Nelson and Winter, 1982). Routines emerge as standard responses to external problems are adopted (Narduzzo, Rocco, and Warglien, 2001). Prior work has noted that routines change slowly in response to external stimuli (Gavetti, Greve, Levinthal, and Ocasio, 2012) and that a slow pace of routine change underpins path-dependencies that can weaken firm performance (Dosi and Kaniovski, 1994; Levinthal and March, 1981), in part because organizations sometimes cease to consider the effect of entrenched routines (Dobrev, Kim, and Carroll, 2003; Dobrev, Kim, and Hannan, 2001; March, Schulz, and Zhou, 2000; Kelly and Amburgey, 1991; Amburgey and Miner, 1992). As yet, however, no microfounded theory explains the stickiness of organizational routines.

3.2.5. Learning

Learning in the Behavioral Theory of the Firm is path-dependent: organizations reinforce behaviors that meet their aspirations, and they adapt behaviors that do not (Gavetti, Greve, Levinthal, and Ocasio, 2012; Levinthal and March, 1993; Cyert and March, 1963). Learning curve models, rooted in individual psychology (Thurstone, 1919), have been extended to the organizational context. A key concept of learning curve theory is "learning by doing" (Maula, Heimeriks, and Keil, 2023). Organizations acquire experiential knowledge through repeated task iterations (Argote, 2012; Fiol and Lyles, 1985; Huber, 1991). Organizations store this knowledge in repositories, transferring it to new activities to influence future task completions based on prior experiences (Daft and Weick, 1984; Huber, 1991; Nelson and Winter, 1982). Studies have explored how organizations codify experiential knowledge, forming routines and capabilities as repositories (Heimeriks, Schijven, and Gates, 2012; Zollo and Winter, 2002). Codification enhances the explicitness and accessibility of tacit knowledge, contributing to organizational performance. Moreover, research has also highlighted the challenges of knowledge transfer through replication (Jensen and Szulanski, 2007).

Behavioral learning theories, specifically performance feedback theory, propose that organizations alter their behavior in response to feedback, with negative feedback prompting change and positive feedback leading to replication (Greve, 2003). This pattern explains the positive organizational experience–performance relationship (Baum, Rowley, Shipilov, and Chuang, 2005). The initial strategies provided by founding teams are important in this account of learning, because they provide the starting point from which adaptation and reinforcement start (Simons and Roberts, 2008; Beckman, 2006). Organizations can learn by introspection from their internal procedures, or from the external environment (Levitt and March, 1988). Recent work has examined organizations' capacity to learn from their competitors: for example, Lee and Pennings (2002) examine learning about organizational structure; Henisz and Delios (2001) learning about location; and Greve (2009) technological learning. But, in practice, while organizations can copy the externally observable characteristics of their peers, it is very hard for them to infer the interdependencies between those characteristics (Zander and Kogut, 1995; Simon, 1962). As a result, some peer-group learning seems very hard (Mueller and Rauh, 2018), and even minor errors can generate significant

penalties (Rivkin, 2000; March, 1991; March, 1981). Notwithstanding these seminal contributions, learning curve studies have focused on learning curve patterns rather than the causal mechanisms underlying these patterns (Maula, Heimeriks, and Keil, 2023).

3.2.6. Inertia

A large literature provides empirical support for the importance of capabilities in explaining incumbent inertia and subsequent failure (Cooper and Schendel, 1976; Henderson and Clark, 1990; Tushman and O'Reilly III, 1996). Established firms face an increased likelihood of failure when confronted with technologies outside their competence domain, demanding mastery of new scientific disciplines (Tushman and Anderson, 2018). The historical trajectory of a firm constrains its learning, primarily guided by localized search processes (March and Simon, 1958; Levitt and March, 1988; Teece, 1988). Additionally, non-technological assets significantly influence technological trajectories, with firms more likely to develop technologies aligned with existing complementary assets crucial for successful technology commercialization (Teece, 1986; Helfat, 1997).

While most of the organizational inertia literature has focused on capabilities, some scholars have focused on the role of cognition in explaining organizational inertia (Garud and Rappa, 1994). A central challenge for organizations facing radical technological discontinuities is the ability to distinguish between changes that require only the development of new from changes that also require the adoption of different strategic beliefs (Maula, Heimeriks, and Keil, 2023). Cognitive representations, rooted in historical experience, play a fundamental role in shaping managerial decisions, resulting in the development of shared beliefs or 'dominant logic' within firms over time (Prahalad and Bettis, 1986). This shared cognition, influenced by firm founders, often impedes adaptability, particularly in dynamically changing environments (Barr, Stimpert, and Huff, 1992; Brown, Crowley, and Elliott, 2020; Tripsas and Gavetti, 2017).

Decision-makers' cognitive representations are frequently rooted in historical experience rather than current environmental knowledge (Kiesler and Sproull, 1982). Notably, senior managers collaborating over time tend to formulate a set of beliefs, often termed 'dominant logic,' shaped by their shared history (Prahalad and Bettis, 1986). These beliefs encompass a collective understand-

ing of relevant competitors (Porac, Thomas, Wilson, Paton, and Kanfer, 1995). The imprint of firm founders on these beliefs is substantial, persisting within the organization long after their departure (Baron, Hannan, and Burton, 1999). In dynamic environments, the influence of the historical context poses challenges for top managers to adapt their mental models swiftly, leading to suboptimal organizational performance (Barr, Stimpert, and Huff, 1992; Brown and Eisenhardt, 1998). Managerial cognition shapes the evolution of capabilities and contributes to organizational inertia; Tripsas and Gavetti (2017) contend that by constraining and guiding search activities linked to technology development, managerial cognition directs the development of new capabilities.

Because organizational feedback is delayed (Rahmandad, 2008) or complex (Denrell, Fang, and Levinthal, 2004; Fang and Levinthal, 2009), organizations often fail to learn from experience. If tasks are causally ambiguous, organizations fail to interpret what has worked and what has not worked, which leads to no or to superstitious learning (Levitt and March, 1988; Zollo, 2009). Given these challenges, studies have advocated for the adoption of deliberate learning efforts by organizations to systematically capture knowledge and cultivate organizational routines (Zollo and Singh, 2004). These intentional learning initiatives aim to unveil cause-and-effect relationships, allowing organizations to extract valuable insights from their experiences. This process involves codifying these insights for subsequent use (Zollo and Winter, 2002). Notably, research in acquisitions (Zollo and Singh, 2004), alliances (Kale and Singh, 2007), and customer relationships (Mayer and Argyres, 2004) underscores the significance of deliberate learning in strengthening the relationship between organizational experience and performance. Taken together, these studies provide robust support for the proposition that deliberate learning efforts strengthen learning by doing (Maula, Heimeriks, and Keil, 2023; Dosi, Pereira, Roventini, and Virgillito, 2017).

Decision-makers, constrained by bounded rationality and information processing limitations (Simon, 1947), may overlook pertinent experiential knowledge due to the application of decision heuristics during complex and time-pressured decision-making scenarios (Eggers and Kaplan, 2013). Melone (1994) observed that executives with financial backgrounds exhibited a tendency to concentrate on financial aspects when assessing potential acquisitions, thereby disregarding prior experience and other strategic factors influencing the desirability of the proposed acquisition. Past

experience can also be intentionally ignored due to organizational power and politics among organizational coalitions (Blackler and McDonald, 2000; Contu and Willmott, 2003; Coopey and Burgoyne, 2000; Cyert and March, 1963). Finally, research exploring heuristics and cognitive biases has also emphasized the crucial role of decision-maker intuition; decision-makers frequently dismiss solutions grounded in organizational experience based on their intuitive or "gut feeling" (Calabretta, Gemser, and Wijnberg, 2017; Khatri and Ng, 2000).

3.3. Language Games: Meaning as Use and Family Resemblances

This section explains Wittgenstein's (1953) claim that "meaning is use." Wittegenstein's account of this idea relies upon the *language games* that run through our social lives; he claims that such games cannot have formal rules and that they are understood in terms of *family resemblances* that are identified and policed by communities. We will use these ideas to understand how organizations categorize environmental cues, and to argue that the organizational meaning of such cues is whatever action they cause.

Wittgenstein's later philosophy was a response to school of thought that held that language must refer to something. For John Locke (1847), the referent was a thought, so that the purpose of language was to communicate the thoughts in the speaker's mind to an interlocutor. In Wittgenstein's early work, the referent of language was a state of affairs in the world (Wittgenstein, 1921), and that work built upon related contributions by Frege (1948) and Russell (1905). But Wittgenstein came to reject this way of approaching language: two people can accomplish things using language without any access to one another's thoughts; it is famously hard to identify the referent in sentences like "the current King of France is bald," and some speech acts are clearly not intended to refer to anything at all.

For Wittgenstein, language makes sense "only in the stream of life" (Malcolm, 2001, 58). This claim rejects the notion that the meaning of a word is whatever image it conjures in the mind; rather, a word means whatever its use attempts to accomplish (PI §§5, 7, 43). That meaning is established in a *language game*: that is, in a social interaction that uses language to get things done. Wittgenstein shows how this works by considering the interpretation of the linguistic sign "five red apples"

in a greengrocer's shop (PI §1). When that sign is uttered, it causes the greengrocer, first, to compare the colour of apples to a standard colour chart (Wittgenstein suggests that the chart is physical, but it need only be commonly understood by the players of the language game), and, second, to incant the sequence "one, two, three, four, five" while transferring apples to a paper bag, before presenting the bag for payment. The purpose of the language game is to transfer applies, and the request for "five red apples" is therefore conative: it acquires its meaning from the effect for which its linguistic community understands it to strive. Similarly, the statement "employee complaint" means whatever it causes a firm's workers to do. Multiple actions may be triggered throughout the firm; each of the associated actors is participating in a language game within which the rule that he or she follows as is experienced as foundational in the same sense that the categorization rule is foundational. The actors concerned need not consider the reasons that an email was categorized as "customer complaint" rather than as "customer suggestion" and, as they play their language game, there is no reason to believe that they necessarily form a moral opinion as to the content of their game.

Wittgenstein's famous private language argument holds that language can only ever be understood in a community. Wittgenstein explains this point with reference to *family resemblances*: by his account, two things should be assigned the same linguistic term if they have sufficient family resemblance. Central to Wittgenstein's account of family resemblance is that it cannot be reduced to a set of codified rules. He explains this statement by asking how we know that something is a "game"; he claims that any attempt to identify an underlying essence from which a definition of "game" can be extracted must fail, because there is no such essence. No single thing links chess, football, and Call of Duty, and yet we believe that all are games. We believe that they are games because the communities within which we live find it useful to classify them in this way. We therefore internalize classification standards that are created, taught, and maintained within the community that we live in. Those classification systems cannot be understood as deriving from rules that we can discover; they are foundational practices that are experienced by individual members of a community and are understood without resource to other explanations (PI §§199, 201-2).

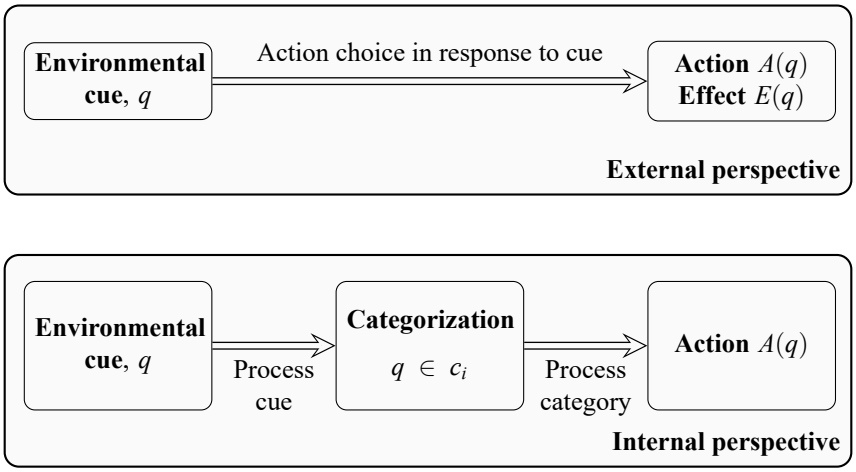


Figure 3.1: **External and internal meanings of environmental cues.**
The Figure illustrates external and internal perspectives on an organization's response to an environmental cue, q. From an external perspective, the organization appears directly to select an action $A(q)$ that is understood externally in terms of its effect $E(q)$, which can be conceptualized as the *external meaning* of q. Internally, the organization makes sense of the cue in two stages. First, it determines which category c_i q belongs to; c_i can be conceptualized as the *categorical meaning* of q. Second, it uses the goal g_i associated with category c_i to select an appropriate action $A(q)$; we conceptualize g_i as the *conative meaning* of c_i. Together, categorical and conative meanings comprise the *internal meaning* of q.

3.4. A Model of Organizational Categorization and Action

3.4.1. External and internal meanings of organizational cues

This section deploys Wittgenstein's insight that "meaning is use" to construct a linguistic model that explains how an organization selects an action in response to an environmental cue, q. Our basic insight is illustrated in Figure 3.1, which shows two ways to understand an organization's response to a cue, q, that it receives from its external environment. The cue might be a customer complaint, a change in the law, or an action by a competitor, for example.

The top part of Figure 3.1 illustrates an external observer's perspective on the organization. From this perspective, the organization can be viewed as responding to a cue by selecting an action $A(q)$, and consequently having an effect $E(q)$ on the external environment. An external observer that creates a given q in order to achieve the effect $E(q)$ experiences the organization as an interlocutor in a language game in exactly the same way that the customer that we studied in the previous section experiences the greengrocer to whom he or she says "five red apples."

We use this analysis to conceptualize the *external meaning* of a cue q to a given organization. The cue has meaning to the extent that the organization is experienced as an actor that is capable of participating in language games. Then, if meaning is use, the external meaning of the cue q is whatever effect $E(q)$ external observers understand that it will achieve when it is presented to the organization.

The difference between the language game in the greengrocer's shop and the language game between the customer and the supermarket chain is that the greengrocer is a single person, while the supermarket is a community of people. We have argued that linguistic understanding for natural, biological, agents is foundational. That means that natural people cannot attempt to resolve misunderstandings in language games with other people by thinking about, and possibly modifying, the biological and physical processes that govern their behaviour in the language game; resolutions of this type of problem come instead from social engagement and discussion. In contrast, an organization is not a natural agent, and so cannot understand a language game as a single person does. The external meaning of an organization's behaviour is an amalgam of the behavior of the natural agents that act for the organization. Hence, if an organizational agent behaves in unexpected ways in its language games, it is possible to analyse factors internal to the agent to understand its behaviour; indeed, it is necessary to do so in order to achieve meaningful change. That analysis explains organizational responses to environmental cues in terms of the foundational linguistic cognition of the natural agents that comprise the organization.

The bottom half of Figure 3.1 presents a two-stage model of the internal processes by which an organization creates the external meaning $E(q)$ of a cue, q. Each of the two stages is accomplished in a separate language game. In the first stage, the organization assigns the cue to a *category* c_i. Categories describe groups of cues that are related by the type of family resemblance that we discuss above. For example, "minor complaint," "competitor announcement," and "employment dispute" are all possible categories. Each category c_i has an associated *goal* g_i, which is an understanding of the effect for which the organization strives in the external environment when a cue is assigned to category c_i.

In the second stage of the organization's internal response to a cue q that was assigned in the first

stage to category c_i, the organization selects an action $A(c_i)$ in pursuit of the goal g_i of category c_i. As discussed below, the action selection could be deliberate or automatic. The action $A(c_i)$ has an effect, which may not precisely reflect the organization's goal g_i; that effect is publicly observable and it constitutes the external meaning $E(q)$ of the cue.

The two stages of the internal perspective correspond to two different types of meaning. The first stage understands a cue q to be a trigger for negotiation over the family resemblances that relate the members of any category in our model. We refer to the category c_i to which q is assigned in this stage as the *categorical meaning* of q. The second stage understands the assignment of q to c_i as the starting point in a separate language game that pursues the goal g_i in the external environment. The goal g_i is a target towards which the organization strives, but that it may never fully accomplish. This type of target is sometimes referred as conative, and we therefore refer to g_i as the *conative meaning* of the category c_i. We present example conative meanings in our section 3.6.2 discussion of Kodak and Fuji. The categorical and conative meanings of a cue together comprise its *internal meaning*.

It is difficult for organizations to assess their performance, because they may never completely meet a goal g_i. In our model, an organization identifies an aspirational level of performance related to the goal, and evaluates its performance against that level. In line with the behavioural theory of the firm, aspirations are established contextually, by examining the organization's historic experience as well as the competitive landscape (Levinthal and Rerup, 2021; Cyert and March, 1963; March, 1962). For convenience, we say that an action $A(q)$ *satisfices* when it meets the aspiration associated with q's category.

3.4.2. Process model

Figure 3.2 presents a process model that shows how an organization constructs, maintains, and revises the private meanings of cues. We consider an established organization that has a standard set of categories $\{c_1, c_2, \ldots, c_n\}$ with associated goals $\{g_1, g_2, \ldots, g_n\}$. In our model, categorical meaning is understood similarly to natural language and, hence, we refer to the organization's categories as its *lexicon*. As indicated in the Figure, the organization's engagements with the external

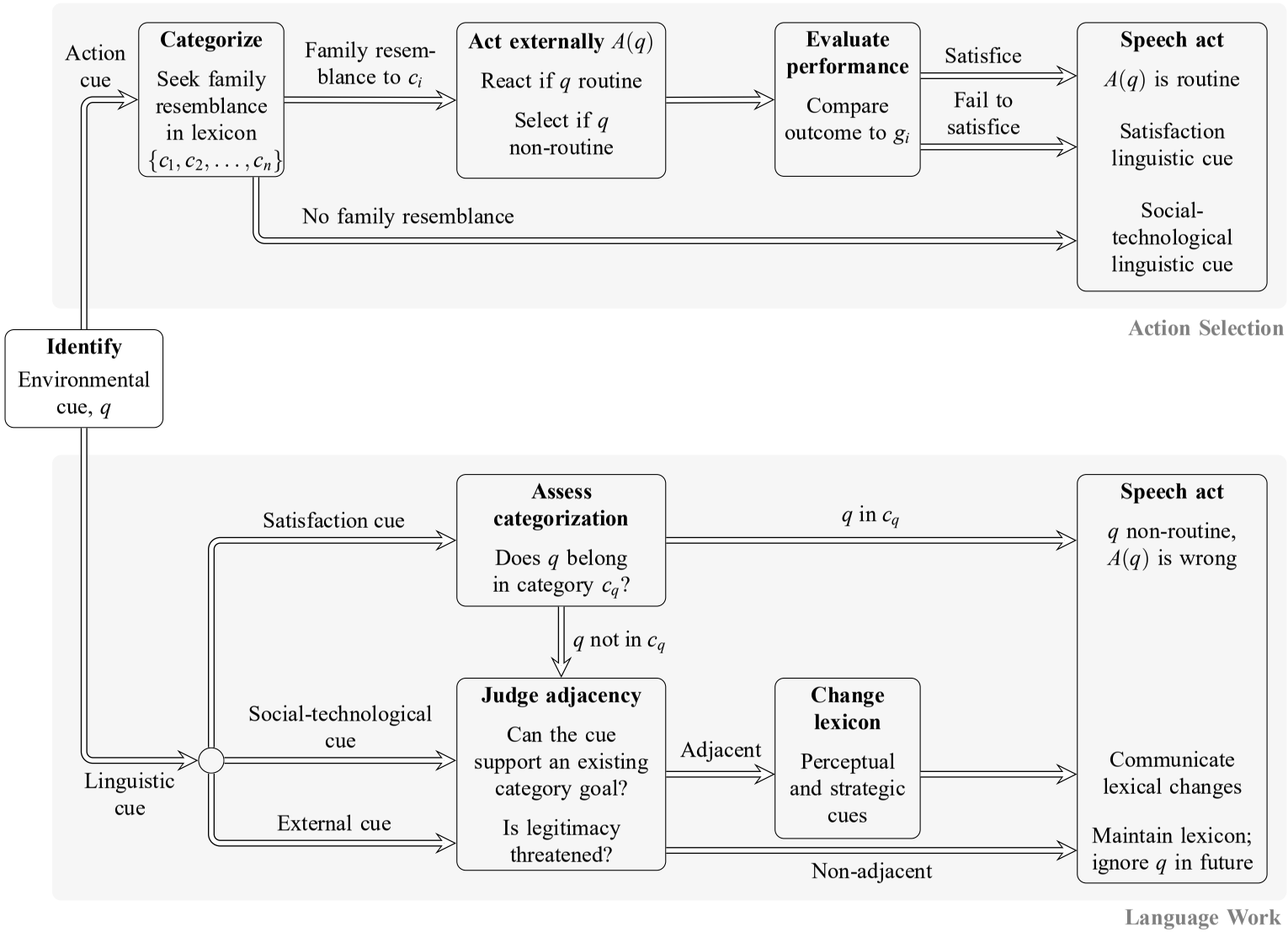


Figure 3.2: **Process Model.**
This process model shows how an organization constructs, maintains, and revises the private meanings of cues. We consider an established organization that has a standard set of categories $\{c_1, c_2, \ldots, c_n\}$ with associated goals $\{g_1, g_2, \ldots, g_n\}$. In our model, categorical meaning is understood similarly to natural language and, hence, we refer to the organization's categories as its *lexicon*. The organization's engagements with the external world have three stages: identification, categorization, and action.

world have three stages: identification, categorization, and action. We start our analysis by examining the identification stage, and we then discuss categorization and action in the respective cases where external cues demand action selection and language work. Throughout our exposition, bold font paragraph names identify boxes in Figure 3.2.

Identify. Our process model starts at the moment when an organization identifies an environmental cue, q. An environmental cue is something external to the organization that demands a response: it might be a demand for higher pay, a complaint about product quality, or an objection to the organization's policy on diversity and inclusion, for example. The organization may be a passive recipient of a cue, as, for example, when it is presented by a prominent external stakeholder like a government or a consumer rights group. But some cues have to be actively sought before they can be identified. Hence, the organization may deliberately scan its environment for cues. The way that it scans, and the cues that it recognized, are informed by the lexicon that it uses to make sense of the world and, in ways that we explain in our discussion of language work below, by its internal policies and past experience. Hence, for example, a financial institution that has no category corresponding to "green finance" may fail to identify cues relating to emerging financing techniques like green bonds.

Environmental cues create two types of stimulus in our model. *Action cues* demand that the organization act on its external environment in some way; the process by which this occurs is illustrated in the region labelled "Action Selection" in Figure 3.2. *Linguistic cues* demand that the organization deliberate over its lexicon; the corresponding organizational processes appear in the "Language Work" region of Figure 3.2. For example, a request for a contractual payment is an action cue, while a demand that the organization change way in which it refers to a minority ethnic group is a linguistic cue.

Action Selection

The "Action selection" region of Figure 3.2 is triggered by action cues. We consider in turn the way that those cues are categorized; the way that goals are identified, and the way that the organization selects actions.

Categorize. As illustrated in Figure 3.2, an organization processes action cues by categorizing them; a cue belongs in a category if the cue bears a sufficient *family resemblance* to other things that belong in the category. Family resemblances are determined by individuals, but they reflect the accumulated choices made by past and present members of the organization and, hence, organizational categorization is path dependent. History matters because the rules for category assignment are linguistic and bottom out in the community activity that governs the assignment. We explain this link by presenting an example categorization and then using Wittgenstein's account of linguistic meaning to explain the nature of the rules that govern categorization.

To illustrate how family resemblances generate categorical meaning, suppose that a firm receives an employee's complaint that she was denied promotion. That complaint could be assigned to category c_1, "employee grievance," or to category c_2, "archival information." One of these categories has the goal that complaints receive a formal response; the other, that complaints be recorded for reporting purposes only. In our model, category selection does not require organizational members to judge whether it would be more normatively appropriate to respond formally to the complaint or to record it. Instead, the organization views a complaint as a grievance if it looks sufficiently like other things that the organization understands to be grievances, and as archival information only if it looks like other things that the organization has treated in the past as archival information. The comparison is made by examining relevant features of the complaint and of members of the respective categories; what "relevant" and "features" refer to is established by within an organizational community.

Note that, by this account, what individuals are *not* doing in the example of the previous paragraph is comparing the employee's complaint to private mental objects that represent prototypical grievances and archival information in order to pick the category whose prototype is, in some precise sense, "closer" to the employee complaint. Indeed, we claim that it is meaningless to talk about this type of mental object: like the connections between the class of "games" studied by Wittgenstein, connections between employee grievances or between archival information are established in a language game that is understood in terms of its effect on the world.

Categorization is a social activity; accordingly, if the appropriate category for unusual or un-

familiar cues is not obvious, then it must be negotiated by members of the relevant organizational community. That negotiation requires the relevant community (1) to determine which features generate family resemblances and (2) to agree which putative family resemblance is strongest. The results of that negotiation could vary between organizations: for example, one firm might understand a "vacation" as involving a total break from email correspondence, while another might understand it to involve continued but intermittent availability.

As environmental cues become more familiar, the organization no longer needs to deliberate over their meaning. When that happens, categorization becomes a natural activity that is best understood as a linguistic phenomenon that, *pace* Wittgenstein, cannot be further explained in terms of deeper algorithmic rules. For familiar cues of this type, deviation from the organization's standard categorization is the same type of error as the misuse of a word in natural language. That is, it feels as unnatural to the organization's members to violate organizational norms when ascribing an environmental cue to the category "vacation" as it would to misuse the word "blue" (say). Standard usage, rather than mental objects or formal rules, drives the communal response to such errors and, hence, is its own justification.

We have referred in our discussion of categorization to the "relevant communities" that perform that categorization. Those communities exist at one of four levels. First, they may be small groups of specialists within the organization so that, for example, the IT department may understand family resemblances that other people do not. Second, parts of the organization may draw upon professional communities of interpretation like those that govern the legal and accounting professions. Third, some classifications may be understood throughout the organization as reflecting common cultural understandings. Fourth, some classifications are societal.

Act externally. When an environmental cue q has identified as belonging to a category c_i on the basis of its family resemblances to other cues that belong to c_i, the organization decides what action to take. In our model, the decision hinges upon whether the organization habitually processes the cue; in case it does, we say that the cue is routine. We consider the routine and non-routine cases in turn.

The organization *reacts* to routine cues. That means, first, that the organization has a standard

response $A(q)$ to the categorization of a routine cue q; and, second, that the individuals who take the standard action for which the cue calls do not stop to consider why they take that action in response to a categorization.

When a non-routine cue q is assigned to a category c_i, the organization *selects* an action $A(q)$. To do so, it first identifies the goal g_i associated with the category and then selects its action accordingly. The goal g_i could be identified in two ways. First, it might be clearly *codified* in a policy that was created by organizational Language Work. But the goal may not be clearly pinned down by such a policy and, over time, the policy may become tacitly understood rather than explicitly codified. Hence, second, the goal may be *inferred* from past actions taken in response to cues in category c_i.

The organization must deliberate over the action that it takes in pursuit of the goal g_i that it identifies for a non-routine cue q. That deliberation could be *rational* or *mimetic*. Rational deliberation occurs when the firm engages in a means-end rational calculation to select an action in pursuit of goal g_i. For example, a firm that identifies a category "customer complaint" might have the goal that it both learns what caused the complaint and also determines whether it ought to alter its processes in response. The organization could attempt to design processes and procedures that accomplish these goals as effectively as possible. Mimetic deliberation occurs when the firm proceeds by analogy with other cues in the same category; doing so may be cheaper and more convenient, but it may also sacrifice some efficiency. Moreover, because the result in mimetic deliberation may deviate from the goal g_i, it can seed routines that, when they are used in goal inference, result in unplanned and path-dependent goal evolution.

Evaluate performance. Recall from section 3.4.1 that organizations assess their actions against aspirational levels of performance; they do so because goals are conative targets towards which organizations strive, but that they may never fully satisfy. Hence, as in section 3.4.2, organizations evaluate their performance by determining whether an action $A(q)$ *satisfices* the relevant aspiration.

Speech act. The last stage of Action Selection is organizational communication; because the associated communication changes the world, we refer to it as a *speech act* (Searle, 2010). We discuss three types of speech acts. The first arises in case an action $A(q)$ satisfices. In that case, the

organization communicates its success internally, and its recognition either renders $A(q)$ a routine action, or reinforces the routines of $A(q)$. This type of speech act therefore alters the way in which Action Selection occurs in future.

The second type of speech act arises in case an action $A(q)$ fails to satisfice. In this case, the organization triggers Language Work. It does so by addressing a linguistic cue to itself that results in deliberation over the categorization of q; we refer to this type of cue as an *satisfaction* linguistic cue.

The third type of speech act occurs when a cue q has insufficient family resemblance to be assigned to a category in the firm's lexicon. This type of cue could be a new technology, or a completely new social practice. The organization responds to this type of cue by triggering Language Work. It does so by addressing a linguistic cue to itself that results in deliberation of the categorization of q; we refer to this type of cue as a *social-technological* linguistic cue.

Language Work

A linguistic cue is an address that demands that the organization perform *Language Work*. By Language Work, we mean the deliberative processes by which an organization assesses the categories in its lexicon and the goals that constitute their internal meanings. Language Work could result in an expansion or contraction of the lexicon; in a change to the goals associated with each category in the lexicon; or in an organizational decision that a cue is a member of some category to whose members it lacks a family resemblance. In order to study Language Work, we discuss the categorization of linguistic cues as "adjacent" and "non-adjacent"; the ways in which lexicons, goals, and category memberships are altered in response to adjacent cues; and the speech acts that organizations take when they perform Language Work.

As indicated in Figure 3.2, Language Work is performed in response to the aspirational and social-technological cues originated during Action Selection, as well as in response to *external* linguistic cues that are originated by an audience external to the organization. Because external parties cannot observe the organization's lexicon or the associated conative meanings, external linguistic cues are expressed in terms of the external meanings assigned to action cues. That is, an

external linguistic cue demands that the external meaning $E(q_a)$ of a given action cue q_a be changed. For example, consider a demand that an organization cease to reinforce old-fashioned stereotypes of women. That demand is an external linguistic cue that refers to a category ("woman") that is recognized by external audiences; note that that category need not be part of the organization's internal lexicon.

Assess categorization. Following the literature on problemistic search, organizations in our model start searching for local solutions to the satisfaction linguistic cues that are raised when an action $A(q)$ fails to satisfice. In our model, a "local" solution is one that does not require the recategorization of q. The first stage in the search triggered by a satisfaction cue is therefore to assess whether the organization might reasonably identify an action that can satisfice for the existing categorization of q. If the organization concludes that this is impossible, then it judges the adjacency of q, as discussed below. If, on the other hand, the organization concludes that another action might satisfice, then it uses a speech act to trigger new Action Selection.

Judge adjacency. Adjacency judgment is the deliberation that occurs in response to an external cue, or when an action cue does not have an organizational categorization. The organization judges whether these cues warrant a linguistic response; when a response is warranted, we say that the cue is *adjacent*. In our model, organizations select actions by asking what their categories are *for*, and then they act in pursuit of those conative meanings. Hence, an organization judges a cue to be adjacent whenever it can help the organization to pursue a conative meaning. We outline the nature of this judgement for satisfaction, social-technological, and external cues in turn.

The organization judges the adjacency of a satisfaction cue raised for an action cue q in category c_i with goal g_i when it concludes that no action $A(q)$ can satisfice for g_i. The adjacency judgment in this case is an assessment of whether q nevertheless has the potential to support an existing conative goal, even though it cannot satisfice that goal directly. For example, donations to universities cannot directly achieve the conative goals of excellence in teaching and research, but they do have the potential to support both goals. Hence, if it does not already have a category, a donation is adjacent to the university.

Social-technological cues are raised for action cues q that bear no family resemblance to an

existing organizational category. Notwithstanding the lack of family resemblance, q could potentially satisfice for an existing conative goal, or could support such a goal indirectly; the cue is adjacent whenever one of these possibilities obtains. For example, the invention of the smartphone was a social-technological shock for taxi cab companies, because it bore no family resemblance to any category in their lexicon. But the smartphone could help them to achieve the conative goal of providing rapid customer transit and, hence, it was an adjacent technology.

Recall that an external linguistic cue is a demand that the public meaning $E(q_a)$ of a given action cue q_a be changed. External linguistic cues are relevant to organizations because they uncover information about the organization's legitimacy and, hence, its capacity to pursue its most foundational conative goals. An external cue is therefore deemed adjacent if it threatens legitimacy in some way. Legitimacy standards are social constructs that are created outside the organization and, hence, are outside the scope of our analysis; they can be viewed as relating to pragmatic, moral, and cognitive considerations (Deephouse, 1996; Suchman, 1995). For example, complaints about so-called sexual harassment do not directly affect an online taxi cab company's capacity to achieve its conative goal of rapid customer transit, but such complaints *do* threaten the company's legitimacy and, hence, they are adjacent.

Change lexicon. Organizations change their lexicons whenever a linguistic cue is deemed adjacent. We distinguish between two types of cue.

First, adjacent *perceptual* linguistic cues require the organization to change the way that it categorizes action cues. Adjacency satisfaction and social-technological cues require the organization deliberately to assign an action cue to a different category or to create a new category for the cue and, hence, are perceptual cues. Adjacent external cues are perceptual when their legitimacy concerns relate to categorization. Table 3.1 presents the possible consequences of organizational deliberation over perceptual cues. There are two possibilities: either the organization creates a new category for the relevant action cue, or it deliberately changes the rules so that the action cue can be assigned to an existing category, even though the organization perceives a weak or non-existent family connection between that category and the action cue.

For example, the MacPherson report (1999) into the response of the London Metropolitan Po-

Linguistic cue type	Revision to lexicon	Speech act
Perceptual	1. New category; or 2. Change category rules.	1. Extend lexicon; or 2. New categorization policy.
Strategic	1. Change category goal; or 2. Split category.	1. Alter action policy; or 2. New categorization policy

Table 3.1: **Responses to linguistic cues.**
This Table summarizes the impact of two linguistic cue types, namely *Perceptual* and *Strategic* on revisions to an organization's lexicon and associated speech acts.

lice Force ("the Met") to the 1993 death by stabbing of the black teenager Steven Lawrence addressed an external linguistic cue to the Met. The Report demanded, inter alia, that the Met stop the disproportionate use of stop-and-search tactics in its dealings with black Londoners. This was a perceptual linguistic cue, because it demanded that black Londoners be perceived in a different way by the Met. The appropriate response from the Met was deliberately to change category rules so that the Met ceased to use race to categorize Londoners.

Second, adjacent *strategic* perceptual cues require the organization to alter the conative meaning that it attaches to the categories in its lexicon. Every strategic cue is external. Table 3.1 identifies two possible responses to this type of cue: the organization can change the conative goal of an entire category, in which case every cue assigned to that category results in a different type of action; or the organization can split the category, so that some action cues induce the organization to pursue the pre-existing conative goal, while others are assigned to the new category and so induce a different action.

For example, Naomi Klein (1999) describes the way that sports shoe manufacturers learned that their shoes were acquiring credibility amongst street gangs who were used as fashion models by suburban youths. The corresponding intelligence can be understood as a strategic linguistic cue demanding that manufacturers change the conative meaning of some shoe categories so that they pursued fashion, rather than sporting, goals. The legitimacy criterion here was pragmatic: treating sporting shoes as fashion items increased manufacturer profits. A sports shoe maker that accepted the legitimacy criterion could respond by splitting a category "sports shoe" into categories "sports shoe for athletics" and "sports shoe for fashion" with different conative goals.

Speech act. As for Action Selection, the last stage of Language Work is an organizational speech act that communicates the results of deliberation and alters future organizational behavior. Three types of speech act are possible. The first occurs after a satisfaction cue that was originated when an action $A(q)$ for an action cue q failed to satisfice. In case the organization concludes that q should not be assigned to a new category, it uses a speech act to inform organization members that q is non-routine and that $A(q)$ is an inappropriate response to q. This speech act changes the Action Selections that occur in response to future cues of type q.

The second type of speech act occurs when the organizational lexicon has changed. In this case, the organization must communicate policies recording changes to categories, to categorization, and to conative goals. Moreover, because the lexical changes cause the organization to act differently in response to some cues, the external meaning of those cues changes, and the organization communicates that fact to external parties, whose engagement with the organization is thereby altered.

The third type of speech act occurs when an organization determines that a linguistic cue is non-adjacent. The organization communicates this fact internally as policy so that, in future, the same linguistic cue is ignored. If the cue was external, then the organization may also respond to the actors who originated it.

3.5. Discontinuous Organizational Change

We have presented a process model in which the categorization of external cues is a distinct activity from action selection; in this section, we show how a partial decoupling of these activities can cause both organizational stability and also discontinuous and unanticipated shifts in organizational behaviour. The reason this happens is that, in our model, categorical and conative meanings are created in distinct language games that are played in distinct communities. As a result, the business of categorization can evolve independently of the practical actions taken in response to that categorization; when that happens, the organization's choices are driven less by means-end rationality than by path-dependent cultural-linguistic factors. In the short run, this effect causes organizational behaviour to remain relatively stable as social and scientific technologies change. But, if the or-

ganization is led by a linguistic cue to revise its category set, then the linguistic understandings that guide its people may lead them to respond to the re-categorization in unexpected ways; such responses can result in significant and unplanned shifts in organizational behaviour.

In short, then, our model explains both inertia and unanticipated behavioural changes as organizational consequences of the decoupling of categorical meaning from conative meaning. We explain how each type of meaning can become decoupled from the other before discussing organizational change in greater detail.

we argue that the the massive upheaval caused by the #MeToo movement can be explained as an unforeseen consequence of relatively small deliberate changes to organizational categorization.

3.5.1. Decoupling categorical meaning

We explain decoupling by examining the creation of categorical and conative meaning in turn. Section 3.4.2 shows how the relevant organizational community uses family resemblances to judge whether a cue belongs in a given category without necessarily considering the effect that the categorization ought to achieve, or that action that it should trigger. Indeed, a community's understanding that a particular cue belongs in a given category is of the same type as the understanding that a particular pigment is properly described as "blue"; as a result, the categorization is generally unexamined and instinctive.

Precisely because categorical meaning comes from family resemblances, it can evolve independently of action choices. To see how, suppose that the organization assigns a novel cue q to an existing category, c. When it does that, the set of category members that the relevant community uses to determine family resemblance to category c changes. That might lead to perceive a stronger family resemblance between some other cue, q', and category c; as a result, q' may be assigned to c for the first time. And that assignment may, in turn, expand the membership of c still further. All of these changes happen without reference to the action $A(c)$ triggered by category c.

This discussion shows that, in our model, categorical meaning is path-dependent. Suppose that new cues q_1 and q_2 bear a weak family resemblance to one another and that the organization perceives each in turn. If q_1 arrives first, then the organization identifies a family resemblance to

category c and then, because q_1 is in category c, q_2 is also assigned to that category when it is perceived. But, if q_2 arrives before q_1 has been categorized, q_2 is perceived as meaningless and, hence, is ignored by the organization. In section 3.6, we show how this observation can affect an organization's response to new technologies.

The U.S. Supreme Court case *Charles E. Moritz v. Commissioner of Internal Revenue* (469 F.2d 466 (1972)) illustrates the evolution of categorical meaning. Charles Moritz argued to the Court that he had been discriminated against on the grounds of his gender; the Court was therefore called upon to assign his case to one of two categories: discriminated, c_d; or not discriminated, c_n. Prior to this case, the Court had never assigned behaviour towards a man to category c_d; that is, family resemblances to c_d appeared to require a female plaintiff. The Court found in favour of Moritz and so concluded that he had been the victim of sexual discrimination. The finding altered the family resemblances that guided categorization in cases of this type, because cases with male plaintiffs could now be assigned to c_d. That happened when the Court found for the male plaintiff in *Weinberger v. Wiesenfeld* (420 U.S. 636 (1975)). And, even when a case was not assigned to category c_d, its categorical meaning changed, because it was now perceived as possible that it could have been categorized in this way.

3.5.2. Decoupling conative meaning

The communities that create conative meanings can operate without considering the family resemblances that underpin categorical meaning. This effect manifests for both the routine categorizations that do not require conscious action choice, and for the non-routine categorizations that do.

The community that selects organizational actions understands that, when a routine cue q belongs to category c_i with conative meaning g_i, it should take action $A(q)$ in pursuit of the goal g_i. That understanding can be maintained without a detailed understanding of the family resemblances that inform the categorical meaning associated with c. Hence, as categorical meanings evolve, the action taken in response to a given categorization need not. As a result, actions that were a means-end rational response to a categorization c may drift away from rationality over time.

An extreme example of this type of drift occurred at the energy company Enron, which was

the seventh-largest publicly traded corporation in the United States when it declared bankruptcy on December 2, 2001 (Senate Subcomittee, 2002, p. 1). At the heart of Enron's failure were a collection of very complex Special Purpose Companies (SPCs) that purported to remove risk from the Enron balance sheet but that did not (Powers, Troubh, and Winokur, 2002). For example, the SPC created to insulate Enron from a drop in the value of its holdings of the New Power Company (NPC) was capitalised with Enron's NPC stock and, hence, did not reduce Enron's exposure at all (Fox, 2003, p. 151). But, within Enron, the SPCs were categorized as bankruptcy-remote by a community that included the accountancy firm Arthur Andersen. As Enron and its accountants started to alter the basis on which they assigned SPCs to this category, its meaning changed. But the actors who relied upon the categorization did not, and were not expected to, understand this change; the result was a slow drift from the means-end rationality that had informed Enron's procedures when they were designed. Ultimately, Enron's communications specialists and lawyers failed to respond to falls in the value of NPC stock and other Enron assets so that Enron failed to reinforce its balance sheet and, in due course, the company failed.

Conscious action choices are triggered by non-routine cues. As discussed in Section 3.4.2, the resultant deliberation could be rational or mimetic; in either case, the action selection creates a new, possibly means-end rational connection between cue and action from which the organization can drift as categorical and conative meanings uncouple in the future.

3.5.3. Organizational change

We now examine our model's implications for the ways that organizational actions change in response to alterations in the external environment. We show that organizational choices can deviate from strict means-end rationality when the processes that create categorical and conative meanings become decoupled.

First, an organization may under react in response to changes in its environment. The reason is that, if the set of categories that the organization uses is fairly stable, it may respond to novel environmental cues by assigning them to existing categories, rather than by creating fresh categories. That is, the effect of a new cue may simply be to change the categorical meaning of a given

category, and that change occurs without any alteration in the conative goal of the category. As a result, the organization may respond to novel cues by pursuing outmoded or inappropriate goals.

Second, adjacent linguistic cues can result in significant unexpected changes to organizational actions. The organization anticipates that some action cues will be recategorised, and so will result in changed organizational actions, when it deliberately changes its lexicon in response to adjacent linguistic cues. But those recategorizations serve to alter perceived family resemblances so that some action cues may be recategorized in unanticipated ways; the result may be sharp, discontinuous and completely unexpected changes to organizational behaviour.

3.6. Case Studies

In this section, we use our model to explain two important, and significantly different, organizational case studies. First, we use the #MeToo movement to illustrate the application of our model to changes in societal understandings of the world. Second, we use our model to shed light upon the very different responses of Kodak and Fuji to the invention of digital photography. Our model explains why, in both cases, organizational responses to change depend upon organizational history and upon the strength of organizational goals in ways that are path-dependent and that are generally impossible for the relevant decision makers to foresee.

3.6.1. The #MeToo movement

Although the activist Tarana Burke had used the phrase “me too” to express her empathy for women of colour who had experienced sexual abuse, the modern #MeToo movement was triggered by a social media post that the actress Alyssa Milano made when commenting about evidence of sexually predatory behaviour by the film producer Harvey Weinstein. The movement highlighted the systemic nature of workplace sexual harassment and of gender discrimination more generally; it resulted in the resignation of several high-profile male executives who were alleged to have a long history of inappropriate sexual behaviour.

In the context of our model, the #MeToo movement can be conceived as a linguistic cue. Be-

haviour that might previously have been treated as "boys being boys" or "locker room talk," and so treated as a minor misdemeanour, was recategorized as "sexual harassment." Behaviours that were categorized as sexual harassment were non-routine action cues that required a consciously selected action. The immediate consequences of this change were as anticipated. Over 200 men lost their jobs in the first year of the movement and almost half of them were replaced by women (Carlesen, 2018). A recent paper shows that many male producers who had worked with Harvey Weinstein in the past hired 40% more female writers after #MeToo. This improvement was not just the result of adding "token" female writers, as the teams sizes did not change (Luo and Zhang, 2021). More generally, sexual harassment at work appears to have dropped in the wake of the #MeToo movement (Johnson, Kidwell, Lowe, and Reckers, 2017); some US states banned non-disclosure agreements relating to sexual harassment; states introduced new protections for workers (Conklin, 2020); sex crime reporting increased by 7% in the months following Milano's first post (Levy and Mattsson, 2020); and gender harassment reporting increased from 76% of women in 2016 to 92% of women in 2018 (Johnson, Keplinger, Kirk, and Barnes, 2019). Moreover, while companies could view sexual harassment largely as an insurable transaction prior to the #MeToo movement, they were forced in the post-#MeToo era to consider the much more harmful reputational effects of sexual harassment.

Some striking evidence has emerged that the #MeToo movement had some unanticipated effects. Some of those effects are consonant with the intentions of the movement: for example, it is possible that the movement's progenitors did not consciously intend that it should end the tipped minimum wage, but serious attempts to make this change appear to have been triggered by the movement. However, many of the #MeToo movement's effects were surely not desired by its founders. For example, Atwater, Tringale, Sturm, Taylor, and Braddy (2019) report that more than 10% of both women and men were less willing after #MeToo to hire attractive women, that 22% of men and 44% of women were more likely to exclude women from social interactions outside the office, and that nearly one in three men were unwilling after #MeToo to have one-to-one meetings with women. Along similar lines, research from LeanIn.org and SurveyMokey revealed that 60% of male managers were uncomfortable participating in common, job-related activities with women,

such as mentoring. Senior-level men are twelve times more likely to hesitate to take one-to-one meetings, nine times more likely to hesitate to travel together for work, and six times more likely to hesitate to have work dinners. These patterns of avoidance are not only harmful to women, but also potentially illegal under Title VII (Conklin, 2020).

Our model can explain the unintended consequences of the last paragraph. The key point is that behaviours that are not in general sexually inappropriate might be categorized with those that are because the two are perceived as bearing a sufficiently strong family resemblance to one another. That perception is more likely to obtain in the wake of a movement that renders sexual harassment salient and, hence, that may serve to alter the dimensions along which resemblances are identified. According to our model, once an organization categorizes a particular behaviour as "sexual harassment," it responds to that behaviour with whatever action it chose when it rationally deliberated over its best response to behaviours that were unambiguously harassment. All of this is understood by the organization's members. A man who believes that a female colleague might perceive a lunch invitation as an attempt to exert coercive power over her might respond by no extending the invitation; and, if he faces the risk that this type of invitation may be miscategorised, then he may try to avoid it entirely by altering his recruitment policies. According to our model, once family resemblances have been established and actions have become standardized, this type of unintended consequence can be addressed only by creating a new linguistic cue.

3.6.2. Responses to new technology: Kodak and Fuji

We now study the responses of Kodak and Fuji to the 1975 invention of digital photography by Steve Sasson, a Kodak engineer. At least ostensibly, both firms were engaged in very similar businesses at that time, and both were in a position to access the new technology. And yet, as witnessed by a voluminous literature of case studies and press commentary, the firms had very different responses to digital photography, which was adopted at Fuji with a degree of enthusiasm that Kodak could not muster. As a result, Fuji remains a market leader in the photography sector, and Kodak has now ceased to trade.

In 1975, Kodak was the market leader in high-quality photographic technology. The firm's

culture was built on organizational myths and stories created during the eighty years of success following its 1892 founding by George Eastman (Lucas and Goh, 2009). Kodak saw itself as both market leader and standard setter in high-quality image production (Gavetti, Henderson, and Giorgi, 2005).

Fuji was founded later than Kodak, in 1934. It operated as a lower-cost operator than Kodak and it never achieved the same level of reputation as Kodak for its films and its chemical development processes. Instead, Fuji's customers were arguably much more concerned with convenience and price than with quality, and the firm developed its products and its market approach accordingly.

We do not believe that we can reverse-engineer the internal workings of companies as they operated forty years ago. But, for the purposes of this analysis, it is reasonable to claim that Kodak and Fuji both recognized a "Photography" category; they may have subdivided photography into subordinate categories, but those are not necessary for our exposition. What is important is the possibility that the different cultural understandings of Kodak and Fuji may have manifested in different conative meanings for the Photography category. We capture the different in approaches of the two firms by positing a goal of "high quality printed images" for Kodak, and of "easily shared images" for Fuji. Our purposes in positing these goals is not to claim that they capture every aspect of the two firms' businesses, but, rather, to demonstrate that this linguistic difference is sufficient in the context of our model to generate a parsimonious explanation for the very divergent experiences of the two firms.

We examine the consequences of two cues for Kodak and Fuji, and we show how the different conative goals of the two companies caused them to act in response to those cues in very different ways. The first cue is the invention of the digital camera; the second is the development of LCD screen technology. We frame our discussion of each cue in terms of our process model.

The digital camera created an action cue q_{dig}. Digital cameras were recognisably cameras and so bore a clear family resemblance to traditional film-based cameras. Hence, in our model, digital cameras were placed in the Photography category by both Kodak and Fuji. Digital cameras were novel and, hence, their emergence comprised a non-routine cue; the firms therefore deliberately selected an action $A(q_{dig})$. For the purposes of this exposition, we assume that $A(q_{dig})$ was some

sort of experimentation.

The results of $A(q_{dig})$ were digital images, which Kodak and Fuji assessed against their respective conative goals for the Photography category. Fuji aimed to create easily shared images. Digital images are indeed easily shared, so that they conceivably concluded that $A(q_{dig})$ satisficed. In contrast, Kodak's conative goal was high quality printed images. Early incarnations of the digital camera generated very poor image quality at a high price and, hence, we claim that Kodak concluded that $A(q_{dig})$ failed to satisfice (see Gavetti, Henderson, and Giorgi's (2005) discussion of Kodak's perception of the price and quality advantages of silver-halide film technology). Kodak therefore generated a new satisfaction linguistic cue.

Following our process model, Kodak's first action in response to the satisfaction cue was to determine whether q_{dig} belonged in the Photography category. We claim that, in light of the very low quality of early digital images, Kodak concluded that it did not. Again following the process model, Kodak then decided whether digital cameras were adjacent: that is, Kodak decided whether they could serve one of the goals attached to an extant category. As the firm that invented digital cameras, Kodak was able to appreciate that, while those cameras could not currently generate high-quality printed images, they might one day come to do so (see Anthony's (2016) discussion of Kodak's acquisition of Ofoto and Avantix) and, hence, we argue that Kodak viewed digital photography as adjacent. In line with the process model, Kodak then created a new category, Digital Images (say), to house digital photography; consistent with this claim, Kodak separated its digital imaging operations from its traditional silver-halide photographic division (Gavetti, Henderson, and Giorgi, 2005). The goal of the new category was to develop the technology with a view to serving the superordinate goal of high-quality image production (Barabba, 2011). In line with our analysis, Kay Whitmore, Kodak CEO from 1990 to 1993, characterized the firm's goal as being to "lead the way in film-based digital imaging;" and George Fisher, Kodak CEO from 1993 to 2000, said that "we will always sell film, paper and chemicals. But in the future, we will let people take pictures and scan them in digital form, and we will make money on the different media (CDs or the Internet, for example) or material for output inkjet paper, thermal papers, and traditional silver halide paper" (Gavetti, Henderson, and Giorgi, 2005).

In summary, by our account, no firm could conceivably have understood what the "optimal" response to digital technology was. Kodak and Fuji responded to digital photography in such different ways because, although the firms had similar technological capabilities, the (linguistic) ways in which they understood those technologies differed. As a result, they envisaged the possibilities of the new technology in different ways; that resulted in divergent linguistic understandings of the world and, ultimately, in different actions.

We now turn to the second cue: that is, to the development of LCD screen technology; once again, we frame our discussion in terms of our process model. LCD screens bear no obvious resemblance to the hand-held devices used to capture photographic images and, hence, we suggest that neither firm would have recognised the screens as belonging in the photography category. Each therefore generated a linguistic cue. The new technology would have been viewed as adjacent and, hence, worthy of organizational attention, if it could support an existing conative goal. Early LCD screens had much lower image quality than could be achieved by photographic paper and, hence, the screens presumably appeared unlikely at the time of their inception to contribute to any of Kodak's conative goal. By our account, then, Kodak classified LCD screens as non-adjacent. The resultant speech acts communicated this policy to the actors who screened environmental cues so that they were much less likely to identify screen-related cues as worthy of deliberation in the future.

In contrast, Fuji viewed LCD screens as adjacent. This happened not because LCD screens could necessarily contribute immediately to any of Fuji's conative goals, but because the technology might help it to do so as digital photography developed. Hence, in line with our model, Fuji identified a "Screen Image" category to which the firm assigned LCD screens, and the new category was assigned a conative goal related to image sharing. That action set the scene for Fuji's huge successes in screen technologies (Raffaelli and Snively, 2020). Note the path-dependent nature of that success: in our model, Fuji recognized screens as adjacent only because they had previously identified digital cameras as adjacent.

3.7. Implications for the Behavioral Theory of the Firm

Our analysis shows that the processes by which language is created, taught, and revised provide a bridge between the cognition and actions of individuals and the mezzo- and macro-level institutions within which those individuals exist. In this section, we show how our framework can shed light on many of the phenomena studied in the Behavioral Theory of the Firm.

3.7.1. Attention directedness

Starting from the linguistic understandings of the firm's members, our theory provides a bottom-up account of the way in which the combined choices of those members cause each of them to experience attention directedness in the institutional setting of a given firm. This approach is in contrast to the top-down approach prevalent (Gavetti, Greve, Levinthal, and Ocasio, 2012) in the literature; it nevertheless shows how attention directedness depends upon the firm's goals, its performance, and its history.

The connection between the organizational and the individual in our model is the set of categories that comprise the firm's lexicon. Those categories are firm-level phenomena, but they exist in the minds of individual members of the firm. In our model, individuals do not understand categories as prototypical mental images, but as linguistic objects whose meaning is established in the language games that govern organizational life. That meaning exists in the rules that establish family resemblances, and in the commonly understood goals associated with categories. This type of socially constructed meaning directs attention in three ways, which we consider in turn.

First, because firm members experience categories as linguistic objects, they naturally understand external cues in terms of their family resemblances to organizational categories, and not in any other way. This form of attention directedness relies upon linguistic codes whose structure depends upon the organization's history.

Second, firm members in our model pay attention to organizational processes only when routine actions fail to satisfice with respect to the conative meaning of the relevant categories. That is, attention is directed towards cues that routine processes cannot deal with once again, those processes

are the path-dependent product of organizational history.

Third, when organizational members are called upon to decide whether a social-technological or a satisfaction cue should trigger organizational actions, they do so by considering its adjacency: that is, by asking whether the cue can support an existing category goal. The organization's attention is therefore directed away from cues that are unrelated to the goals that it has accumulated in the past. As discussed in section 3.6.2, this type of attention direction provides a possible explanation for Kodak's failure.

3.7.2. Goals

Our framework provides a fresh perspective on the fact that multiple organizational goals can co-exist in the Behavioral Theory of the Firm. Goals in our model are organization-level constructs that organizational members experience as the (conative) meaning of categories. An organizational member who understands external cues in terms of family resemblances need not experience dissonance when different categories generate very different actions. But it is impossible from the external perspective to identify the internal categorical meanings of cues and the conative meanings assigned to those cues. Hence, from an external perspective, two cues that appear related may generate actions that are hard-to-reconcile. It follows that an external observer may identify ambiguity of meaning that is not experienced by organizational members.

Our model also yields an account of the way that goals evolve. By our account, a cue that cannot be categorized is acted upon by the organization only when it is adjacent: equivalently, when the cue may enable the pursuit of an existing conative goal. Goals are therefore established in a path-dependent way, once established, they alter the organization's adjacencies and so affect its future evolution.

3.7.3. Problemistic search

Our language-based setup gives a natural account of problemistic search that conceptualises the local solutions from which such searches start. As illustrated in Figure 3.2, search is triggered in

our model by the Satisfaction linguistic cue that the organization raises whenever an action $A(q_i)$ taken in response to a cue q_i fails to satisfice with respect to the goal g_i of the category C_i to which q_i belongs. Then, as illustrated in the bottom half of the picture, the organization first assesses whether or not q_i belongs in category C_i. If it does, the organization uses a speech act to instigate an attempt to find a better way to pursue goal g_i; because the goal and the category remain unchanged, this attempt can be understood as being local.

In case the local search for a new action in pursuit of the existing goal g_i fails, the organization judges the adjacency of the cue q. Adjacency is another form of localness, in that it is defined with reference to the organization's existing categories and existing goals. As indicated at the bottom of Figure 3.2, a cue that is not adjacent is discarded so that, in our framework, even organizations that jump to a new practice jump to one that is adjacent to their existing goals and practices.

3.7.4. Routines

Our model yields a micro-level explanation for routines, and it also explains their relatively long-term stability.

A routine is established in our model whenever an action $A(q)$ taken in response to an action cue q satisfices with respect to the conative meaning of the category to which q belongs. In that case, $A(q)$ is adopted as the standard response to q and it is maintained without reflection for as long as the action continues to satisfice. Routines are long-lived when satisficing is also long-lived; one reason that this occurs in our model is that, as discussed in section 3.5.2, conative meaning can become decoupled from categorical meaning so that a category's goals, and, hence, the corresponding satisficing conditions, can remain stable even when that category's composition changes dramatically in response to social or technological factors.

3.7.5. Learning

Our model explains organizational learning as a micro-level phenomenon. Organizational cognition is manifested in our model in the family resemblances that assign categorical meaning to

cues, and in the conative meanings assigned to categories; learning in our model occurs when these markers of organizational cognition change. Learning is a micro-level phenomenon because categorization rules exist in individuals' minds. Those rules are nevertheless created and maintained in an organizational language game; categorization therefore relies upon organizational languages that are passed from generation to generation and, hence, can outlive the tenure of any organizational member.

Organizational learning occurs in our model in case organizational lexicons evolve after actions fail to satisfice with respect to the conative meaning of some category. This observation has two implications. First, conative meanings are essential to the feedback loops that trigger learning when performance is interpreted. Second, because conative meanings may be decoupled from categorical meanings (see sections 3.5.2 and 3.7.4), the satisficing failures required to trigger organizational learning in response to environmental change may occur some time after the relevant changes have occurred.

This framework helps to explain why organizations sometimes struggle to learn from their competitors. When an organization observers its competitors' actions, it does so from the external perspective of Figure 3.1. That is, while organizations can observe their competitors' cue/action pairs, they cannot observe the rules that govern family resemblances in competitors, nor the conative meanings that drive decision making and generate performance feedback. Consequently, it is very hard in all but the most straightforward cases for a firm to mimic its competitors' successes.

3.7.6. Inertia and change

In our process model, the categorization of external cues is distinct from action selection. We argued that this can lead to a partial decoupling of these activities which can both explain organizational inertia and discontinuous change. Recall that categorical and conative meanings are created in distinct language games that are often played by separate communities; one community uses family resemblances to assess whether a cue belongs in a given category without considering the action that this categorization ought to trigger. The two processes can therefore evolve independently of the actions taken in response to that categorization. When this happens, an organization's choices

are driven less by traditional means-end rationality and more by path-dependent linguistic factors. This can both explain organizational inertia and discontinuous change. We discuss both in turn.

The decoupling of categorical and conative meanings can lead organizations to behave in a stable manner, even as social and technological cues external to the organization change. Nonetheless, if the new cues cause the organization to revise its category set, then the change in family resemblances that guide people may lead them to respond to the recategorization in dramatically different ways. First, an organization can under react to environmental changes. If the pre-existing categories used by the organization are relatively stable, then it can respond to new environmental cues by simply assigning them to existing categories, rather than by creating new ones. This would only lead the categorical meaning associated with a given category to change without affecting its conative goal. Consequently, the organization may address new cues by pursuing outdated goals. Second, our model can also explain discontinuous organizational change. If organizational communities anticipate action cues to be recategorized in response to adjacent linguistic cues, then they will deliberately change their lexicon in response to those cues. This, in turn, will change associated family resemblances and lead action cues to be recategorized in unpredictable ways.

3.8. Proposed Empirical Tests

This section outlines a way to test our model.

3.8.1. Hypothesis development and testing algorithm

Hypothesis 1

Our model shows how organizations respond to external cues. It yields the following explanation for situations in which one organization (A) responds to a cue (q) by continuing with business as usual and another (B) that operates in the same markets responds to the same cue by altering the way that it does business. According to the model, A's response does not change because it views q as belonging to an existing organizational category, c, and B's response changes because it finds no family resemblance between q and any of its existing categories.

We therefore have the following hypothesis:

Hypothesis 3.1. *Suppose that two organizations in the same market receive the same external cue. If one of them responds by continuing business as usual and the other alters its business model, then they must be using different sets of categories to assess external cues.*

In order to test this hypothesis, a researcher would need to:

1. Identify a cue q that is relevant to a particular market, M.
2. Identify companies A and B operating in market M, with A responding to q with business as usual and B responding by changing its business model.
3. Identify the categories c_A and c_B that A and B use to assess environmental cues.

Our model implies a null hypothesis that c_A and c_B are different and, in particular, that q bears a family resemblance to an element of c_A and not to c_B. The model should be rejected in case (i) q bears a family resemblance to elements of both c_A and c_B, or (ii) q bears no family resemblance to any element of c_A or c_B; or (iii) q bears a family resemblance to elements of c_B but not c_A.

Hypothesis 2

In our model, organizational inertia results from the decoupling of external cue categorization and action selection. We argue that this process can explain why legacy organizations react differently to environmental shocks than new organizations. Consider an old legacy organization (L) and a new start-up (S) reacting to a technological cue (q_t). If pre-existing categories are stable, then legacy organization (L) can assign new environmental cues to existing categories rather than create new categories. This would only lead the categorical meaning associated with a given category to change without affecting its conative goal. Consequently, legacy organization (L) may address new cues by pursuing outdated goals. In contrast, the pre-existing categories of the new start-up (S) would have more limited family resemblances and would therefore be more likely to create new categories in response environmental shocks. The probability of categorization and action selection becoming decoupled would be lower. This generates the following hypothesis:

Hypothesis 3.2. *Suppose that two organizations, old legacy organization (L) and a new start-up (S), in the same market receive a new cue (q) that is relevant to their business proposition. Suppose also that L responds to q using an old process and S creates a new process. In that case, then L is assigning q to an existing category in the pursuit of an existing goal whereas S is assigning q to a new category in the pursuit of a new goal.*

In order to test this hypothesis, a researcher would need to:

1. Identify a new cue q that is relevant to a particular market, M.
2. Identify companies L and S operating in market M, where L is an old legacy company and S is a new start-up and L took longer than S to respond to q.
3. Identify which categories,c_L and c_S, the organizations assigned q to and which goals, g_L and g_S, were associated with the categories.

Our model implies a null hypothesis that c_L and c_S are different and, specifically, that c_L was altered to absorb q in the pursuit of g_L whereas c_S was split into a new category with a new associated conative goal g_S. The model should be rejected if they have the same categories or the same associated goals.

Hypothesis 3

We argued that adjacent linguistic cues can result in significant unexpected changes to organizational actions. If a legacy organization (L) anticipates that some action cues will be recategorised, and so will result in changed organizational actions, then it deliberately changes its lexicon in response to adjacent linguistic cues. But those recategorizations serve to alter perceived family resemblances so that some action cues may be recategorized in unanticipated ways; the result may be sharp, discontinuous and completely unexpected changes to organizational behaviour.

Hypothesis 3.3. *If a legacy organization L suddenly responds in a new way to an old cue, then it must be because shifts in perceived family resemblances led to action cues being recategorized in unanticipated ways.*

In order to test this hypothesis, a researcher would need to:

1. Identify a cue q that triggered a sudden change in a legacy organization L.
2. Identify how organization L altered the family resemblances of the relevant category, c_L, in response to q.
3. Identify how action cues were changed in response to this recategorization.

Our model implies a null hypothesis that the sudden change in the legacy organization's behaviour can be explained by consequences that altering family resemblances of the relevant category has on its action cues. The model should be rejected if the discontinuous organizational change was not triggered by a change in the family resemblances of the relevant category in organization (L).

3.8.2. Operationalizing the test

Our model is particularly well suited for comparative case studies. It can be used to understand why organizations react differently to certain technological or socio-cultural shocks by looking at their internal categorization process and associated goals. In order to run the test of section 3.8.1, we need to get inside the firm and identify (i) cues, (ii), categories, (iii) goals, and (iv) actions. We briefly explore how this could be accomplished using qualitative and quantitative methods using the move towards net zero as an illustrative example.

1. *Cues:* In order to assess how internal categorization processes and associated goals impact corporate action, we start by identifying external cues that affected multiple firms heterogeneously. In the case of net zero, a suitable cue might be Larry Fink's letter to CEOs–highlighting that risks associated with climate change will compromise returns without reallocation of capital–or the publication of the Paris Agreement goals. Both cues triggered very different reactions by organizations: some proceeded to set very ambitious decarbonization targets whereas others remained wholly silent on the matter.
2. *Categories:* We then need to establish the category that the organization associated with that cue. This analysis could be done qualitatively through interviews to understand workflows

within the organization that link the external cue to internal organizational processes. It could also be done quantitatively using word embeddings on e-mail corpuses and archival internal communications to understand how the cue was first internalized by the organization, who identified it, and how this interest either grew or was rejected. For example, in response to decarbonization targets set forth in the Paris Agreement, Microsoft instituted an internal carbon tax. Any corporate policy instituted by the organization now needs to consider the costs of negative environmental externalities as part of its bottom line. The company subsequently hired a team of over 200 climate specialists to direct this internal carbon tax budget towards different carbon offsetting projects. Every year, they make an open procurement announcement where project developers can submit bids for contracts.

3. *Goals:* We would also need to identify the goal associated with the relevant category. This could be done by looking at internal performance evaluation documentation such as annual or quarterly targets as well as individual or departmental KPIs. Adjacency judgments in the model derive from this goal. We would need to ask whether pursuing actions related to that cue could support this goal or whether its legitimacy would be threatened if the organization failed to react. In the case of Microsoft, the internal carbon tax was deemed adjacent to its conative goal of "empowering every person and every organization on the planet to achieve more".
4. *Actions:* Finally, we would evaluate how these internal categorization processes resulted in external actions by analysing press releases, quarterly earnings calls, or external communications that outline a change in policy. Moreover, we would also analyse whether these internal categorization processes resulted in any internal changes in organizational processes or procedures.

3.9. Conclusion

This article presents linguistic foundations for the Behavioral Theory of the Firm. It relies upon a Wittgensteinian account of meaning as inhering in the uses to which a community puts language,

and in the common and foundational understandings that they create to establish family resemblances. We use these ideas to show how organizations process external cues by categorizing them and then selecting actions that reflect the use-meanings that organizations assign to their categories. This conative understanding of meaning as derived from purpose generates important path dependencies because, in our model, organizations make sense of, and respond to technologies in light of their understanding of what the relevant categories are *for*. Our linguistic approach yields explanations for many central features of the Behavioural Theory of the Firm. In particular, attention directedness, problemistic search, the importance of learned organizational routines, and the epistemic problems that firms encounter when they attempt to learn from their peers all emerge naturally within our framework.

4

Vocabularies, Attention, and Imprinting in Organizations: An Experimental Approach

(Joint with Alan D. Morrison[1])

Abstract

We study how vocabularies microfound organizational imprinting. Once imprinted, they are long-lived and affect the ways in which organizational members perceive their world and, hence, upon their actions. We present a process model within which organizations converge on specific vocabularies that enable them to perform tasks efficiently and, in turn, those vocabularies support commonly understood categories. We study these phenomena in an experimental setting within which participants communicate orally to perform an image recognition task. We perform Natural Language Processing (NLP) on transcripts of the subjects' speech in order to quantify the formation, transmission and persistence of organizational vocabularies. When subjects leave the experiment, they perform a basic memory task that measures the way that their attention was directed during the experiment. Our set-up therefore facilitates an analysis of the relationship between strategic vocabularies and organizational imprinting. Our work yields five results. vocabularies are imprinted on an organization by the repeated iteration of tasks by its members. Second, organizations can operate more efficiently once a vocabulary has been imprinted. Third, imprinting is profoundly *organizational*: that is, holding personnel, processes, categories, and goals fixed, organizational identity affects the vocabularies that people use to perform tasks. Fourth, the likelihood that a word in a vocabulary affects a organization member's cognition of their environment is determined by the way that the member encounters the vocabulary. Finally, the vocabularies that founders imprint on their organizations have very long-lived cognitive effects on organizational members.

[1]The authors thank Paul Gouvard, Michael Jensen, Marc Ventresca, and Eero Vaara for their comments on an earlier draft, as well as audiences at the Berkeley Culture Conference 2023; the Academy of Management Conference 2022 and 2023; the European Group for Organizational Studies Conference 2020, 2021 and 2023; and at brown-bag seminars at Oxford.

4.1. Introduction

Organizations are devices for cooperation and cooperation is impossible without language. While these facts have always been at least implicitly recognized in organization theory, it has taken a more explicitly linguistic turn of late (Ocasio, Laamanen, and Vaara, 2018). A particularly fruitful approach to this work focuses on the role of organizational vocabularies. A vocabulary is a system of cultural categories (Loewenstein, Ocasio, and Jones, 2012). Culture has two roles to play in this statement: it associates a category with a word or another symbol, and it also creates a common understanding of the constituents of that category. Vocabularies therefore identify and create categories that structure group cognition and so inform logics of action. This chapter examines the ways in which vocabularies become imprinted on organizations. It shows that, once imprinted, an organizational vocabulary has a long-lived effect upon the ways in which organizational members perceive their world and, hence, upon their actions.

Empirical research on organizational vocabularies faces a number of challenges. First, it is difficult in live organizations to characterise the precise structure of a vocabulary: that is, to determine the word frequencies, the word-to-word relationships, and the word-to-example relationships that delineate the vocabulary's practical use. Second, the researcher can never completely disentangle the effects of vocabularies from those of social signals gleaned from organizational actors' past experiences. Third, it is usually impossible to study the formation and evolution of a vocabulary from the moment an organization is created. Fourth, causal relationships are very hard-to-establish for organizational vocabularies. Vocabularies identify particular categories as important; an organization's categories shape its cognition of the world; cognition underpins organizational practices; practices justify and structure organizational actions; and, in turn, actions, practices, and cognitions give rise to vocabularies. This type of complexity renders it extremely hard in real-world settings to isolate the effects that vocabularies have upon attention, cognition, and action. Finally, the relationship between staff turnover and the creation and transmission of organizational vocabulary is also bedevilled with endogeneity problems.

This chapter addresses these challenges by studying in an experimental setting the ways in which vocabularies are imprinted on organizations. Our experimental subjects communicate orally

in order to perform an image recognition task in a hierarchical organization. The subjects develop strategic vocabularies that enable them to perform their tasks efficiently, and those vocabularies become imprinted on the organization as its staff turns over in a predictable way. We use transcripts of the subjects' speech to perform Natural Language Processing (NLP) so as to quantify the formation, evolution, and transmission of organizational vocabularies. When subjects leave the experiment, they perform a basic memory task that measures the way that their attention was directed during the experiment and, hence how they understand their environment. Both the organization's vocabulary and this measure of attention remain stable after the founders leave the organization: that is, as vocabularies are imprinted on the organization, so, too, is a particular understanding of the world. Our set-up enables us to analyse the relationship between these types of imprinting.

By design, organizational categories, practices, and goals are fixed in our experimental setting; staff turnover follows a mechanical rule; and experimental subjects encounter one another as strangers. Our set-up therefore allows us to address many of the problems that researchers face in real-world studies of vocabularies and imprinting. In particular, we are able to isolate the way that vocabularies are imprinted on organizations and how, by directing the attention of organizational members, imprinted vocabularies lead in turn to imprinted cognitive frameworks. Moreover, we are able to place the same people in precisely the same situation in two different experimental organizations and hence, to study the relationship between collective-level cognition and situation-level attention.

Because all of our experimental organizations share the same processes, categories, and goals, we are able in our set-up cleanly to address some important questions about organizational vocabularies. First, we are able in our set-up to measure the way in which organizational members direct their attention. Second, we can investigate the longevity of organizational vocabularies and, in particular, we can examine the effects that founders have upon long-run organizational cognition. Third, we can study the factors that cause an organization to change its vocabularies, and, in turn, we can study the effects of such a change upon organizational attention. Finally, we can examine the effect of changing collective level facts about an organization while holding everything else about the organization, including the identities and roles of its members, constant.

The chapter yields five main results. First, vocabularies are imprinted on an organization by the repeated iteration of tasks by its members. Second, imprinting a vocabulary on an organization causes it to operate more efficiently. Third, imprinting is profoundly *organizational*: that is, holding personnel, processes, categories, and goals fixed, organizational identity affects the vocabularies that people use to perform tasks. Fourth, the likelihood that a word in a vocabulary affects a organization member's cognition of their environment is determined by the way that the member encounters the vocabulary. Finally, the vocabularies that founders imprint on their organizations have very long-lived cognitive effects on organizational members.

4.2. Literature Review

This section argues that organizational cognition is directed by vocabularies that, once imprinted, have a long-lived effect upon the ways in which organizational members perceive their environment and, therefore, upon their actions. We start by claiming that vocabularies are systems of codes that serve as community-specific category markers linking stimuli to labels. We then posit that vocabularies direct the attention of organizational members to specific features of the world around them and can serve as a micro-foundation of organizational imprinting.

4.2.1. Vocabularies as code schemas

A rich literature on culture in strategy was inspired by the toolkit view of culture in sociology, which conceptualizes organizational cultures as sets of "components that are used to construct strategies of action" (Swidler, 1986, p.273). Rather than focusing on cultural alignment through shared values, this stream of research focuses on situations where goals are misaligned and decision makers deploy a set of cultural artifacts or frames to steer the attention, perfection and ideation of collectives according to their own interests (Kaplan, 2011; Rindova, Dalpiaz, and Ravasi, 2011; Martins, Rindova, and Greenbaum, 2015; Weber and Camerer, 2003; Dalpiaz and Di Stefano, 2018; Sinha, Jaskiewicz, Gibb, and Combs, 2020).

The code-based view of culture was born from this theoretical tradition (Kocak and Puranam,

2023). In this context, a code is a conventional linguistic expression that could comprise one token (e.g., "dog") or more (e.g., "coffee shop"). A code serves as a marker for a community-specific category of things in the world, linking stimuli (comprising objects, events, perceptions, and actions) to labels (encompassing words, images, icons, or other symbolic forms) (Arrow, 1974; Cremer, Garicano, and Prat, 2007). Vocabularies, in turn, are systems of codes and their meanings. We learn codes by doing things in communities. Learning a word is therefore a practical and social matter; it requires us to master the associated categories, and the way that it is invoked in our day-to-day doings (Mills, 1940; Mills, 1939; Burke, 1937; Burke, 1934). That allows us more effectively to coordinate our activities (Cremer, Garicano, and Prat, 2007; Weber and Camerer, 2003). They are shaped through personal experiences, observational learning, and social transmission. Regardless of their acquisition method, the similarity of codes within groups and divergence across groups is expected due to the varying nature of influencing experiences. If my community views McDonald's as a coffee shop, then it *is* a coffee shop; in another community, it might not be.

Moreover, when a group learns a vocabulary and its associated ways of acting, it acquires common ground that serves to direct its members attention (Loewenstein, Ocasio, and Jones, 2012). An individual may also have perceptions or expectations about the codes that others use (Kocak and Puranam, 2023). A communication code is in the common ground between two people if both use it and if each expects the other to use it as well (Clark, 1996). Expectations about codes can also create convergent behavior, especially when coordination (or its failure) is critical to outcomes. Fussell and Krauss (1992) present an experiment that demonstrates how expectations related to codes enable effective anticipation of others' actions and coordination of communicative behavior. In this experiment, pedestrians in Boston were approached for directions to a landmark, with the inquiry made either in a Boston accent or a rural Missouri accent. Responses to Boston-accented inquiries were concise and practical, tailored to someone familiar with the Boston area, while responses to rural Missouri-accented inquiries were more elaborate, citing numerous landmarks. The compression of communication influenced by the (expected) shared understanding of causal codes has also been observed in the context of distributed software development (Srikanth and Puranam, 2014).

A code that has high similarity among members of a group is called a *consensus*. An individual

perceives a consensus when their primary expectation is that a group shares a similar code (Zou, Tam, Morris, Lee, Lau, and Chiu, 2009). For codes to drive individuals towards the same action, they must be both similar and crisp (Kocak and Puranam, 2023). Identical yet ambiguous codes may result in entirely divergent actions. Consequently, shared codes are defined as encompassing similar and precise associations. Expanding on this notion, cultural strength pertains to the presence of highly similar and precise codes within a group (Chatman, Caldwell, O'Reilly, and Doerr, 2014). The diversity of codes within a group can be gauged by measuring dissimilarities in codes (in usage or expectation) while accounting for their crispness (Marchetti and Puranam, 2022). Disparities in codes between groups can similarly be assessed across code types. The assembly of codes utilized or expected within a specific group may be encapsulated as a cultural configuration (Chatman, Caldwell, O'Reilly, and Doerr, 2014).

Kocak and Puranam (2023) argues that this code-based view of organizational culture provides a link between formal theory as well as empirical methods. Several existing formal models of culture already use codes or vocabularies. The extensive body of literature on Lewis signalling games and their variations takes lexical codes as a central model object (Spike, Stadler, Kirby, and Smith, 2017). Models addressing the imitation and diffusion of attitudes similarly rely on code-like objects to capture beliefs and associations (Goldberg and Stein, 2018). Further models have explored how the learning of causal codes can undergo change through sorting, socialization and sensemaking (Kocak and Puranam, 2018). Similar opportunities for leveraging insights exist in empirical investigations of culture using Natural Language Processing (NLP) which use a code-based approach to studying the impact of cultural strength (Marchetti and Puranam, 2022), heterogeneity (Corritore, Goldberg, and Srivastava, 2020), and alignment (Srivastava, Goldberg, Manian, and Potts, 2018).

4.2.2. Attention

Attention matters in organizations because, if an organization is successfully to coordinate the actions of its members, it must direct their attention appropriately. When organizations direct attention, they affect which their members experience and reason about the external environment

(Cyert and March, 1963; Simon, 1957; Simon, 1947); in turn, they influence the issues that actors engage with, the problems that they uncover, and the answers that they find (Ocasio, 1997; Ocasio, Rhee, and Milner, 2020). When attention is directed, it can generate organizational blind spots and, hence, undermine organizational learnings and cause stasis (Levinthal and March, 1993).

4.2.3. Imprinting

The seminal Stinchcombe (1965) posited that organizations absorb elements of their founding environment and these elements persist beyond the founding phase. Building on this idea, subsequent research has demonstrated the enduring influence of founding context on various phenomena, including managerial structure (Baron, Hannan, and Burton, 1999), the configuration of inter-organizational networks (Marquis, 2003), and survival rates (Romanelli, 1989).

Neoinstitutional theory, traditionally centered on organizational-level historical research, appears well-positioned for investigating the imprinting process (Marquis and Tilcsik, 2013). However, its historical focus has primarily centered on aspects of organizational life other than founding events (Johnson, 2007). he initial reluctance to explore entrepreneurial roles in founding events was rooted in a neoinstitutional critique of interest-based theories of organizational behavior, emphasizing stability over change (Dobbin, 1994). Nevertheless, the 1990s witnessed a shift as neoinstitutionalists developed theories of agency compatible with institutional analysis, delving into empirical studies exploring the roles of social actors, including entrepreneurs, in shaping and transforming organizations (Fligstein, 1996; Dobbin and Dowd, 1997).

While these studies do not explicitly address organizational imprinting, their advances regarding the nature of entrepreneurship and founders brought about a renewed interest in the imprinting phenomenon itself (Johnson, 2007). Since entrepreneurship is inherently a cultural process, an account of organizational imprinting calls for an assessment of the repertoires influencing founders as they establish new organizations. Persistence is then attributed to the process of institutionalization, wherein stable social arrangements emerge and endure through self-activating social processes, influencing goals, rules, coordination mechanisms, and communication channels within organizations (DiMaggio and Powell, 1983). This institutionalization perspective underscores how

organizations create enduring structures that become ingrained and valued beyond immediate task requirements (Selznick, 1957).The concept of imprinting has garnered attention across diverse fields, spanning organizational ecology (Carroll and Hannan, 1989; Swaminathan, 1996), institutional theory (Johnson, 2007; Marquis and Huang, 2010), network analysis (McEvily, Jaffee, and Tortoriello, 2012), and career research (Higgins, 2005). While much of the literature, influenced by Stinchcombe's insights, has predominantly focused on organizations as the primary unit of analysis, recent years have witnessed the emergence of imprinting research extending to meso and micro levels of analysis (Marquis and Tilcsik, 2013).

We discuss some relevant insights from the literature on (i) career and organizational imprints on individuals and (ii) individual imprints on organizations. Note that there is a rich scholarship on institutional imprinting, but a comprehensive review of the imprinting literature is beyond the scope of this paper (Marquis and Tilcsik, 2013, see).

Organizational and career imprints on individuals

Research on imprinting at the individual level has explored the lasting effects of institutional factors, specifically the norms and cognitive schemas associated with organizational cultures. Organizational culture significantly shapes the human and social capital, assumptions, beliefs, and values of early-career managers, forming a lasting career imprint (Higgins, 2005; Dokko, Wilk, and Rothbard, 2009). Individuals carry these imprints across organizational boundaries, impacting performance and career trajectories. (Dokko, Wilk, and Rothbard, 2009) posits that imprinted cognitive and behavioral patterns act as baggage when individuals transition to a substantially different organizational context, hindering responsiveness and reflection.

The influential power of individual-level imprints is demonstrated in studies by Higgins (2005) and Burton, Sørensen, and Beckman (2002). Higgins highlights how Baxter International, due to imprints on employee careers, produced a disproportionate number of leaders in the biotech industry. Imprints from early career experiences at Baxter in the 1970s influenced line management skills and a bottom-line orientation, aligning well with the entrepreneurial demands of the emerging biotech industry in the early 1980s. In contrast, early career experiences at Abbott Laboratories

during the same period fostered a different imprint, less suited to the evolving biotech sector. Notably, individual-level imprints, reflecting core aspects of organizational culture, extend beyond shaping individual cognition and behavior to significantly impact organizational outcomes.

However, a gap exists in understanding the influence of multiple coexisting imprints, as research has yet to explore how imprints from different environmental forces and sensitive periods interact. The concept of secondhand imprinting introduces the idea of social transmission of imprints, where actors take on aspects of imprints from others (Tilcsik, 2014). This can occur when newcomers imitate veteran colleagues, indirectly learning norms and routines that reflect their experienced coworkers' career imprints. Similarly, organizations may emulate the structures and practices of older organizations, incorporating imprinted patterns from earlier eras (secondhand imprinting). Understanding the coexistence and interplay of intersecting imprints within an actor, arising from different environmental forces, sensitive periods, or secondhand imprinting, offers a promising avenue for future research across all levels of analysis.

Imprints of Individuals on Organizations.

Research on imprinting emphasizes the enduring impact of individual founders on organizations. Qualitative approaches have illuminated how founder imprints persist over time. Kimberly (1979) longitudinal study of a medical school revealed that a founder's personality and initial decisions exerted a lasting effect on subsequent organizational strategies. In the context of a French computer firm, Kimberly and Bouchikhi (1995) described the founder imprinting process, emphasizing the significance of early decisions in shaping the organization's trajectory. Johnson (2007) case study of the Paris Opera underscored the role of entrepreneurs in selecting and incorporating historically specific elements, contributing to an organization's fundamental features for decades or even centuries.

Quantitative research has further substantiated the influence of founder imprints. In Silicon Valley high-tech start-ups, (Baron, Hannan, and Burton, 1999) measured founders' mental models and initial decisions, revealing persistent effects on organizational trajectories even after the founder's departure. Founder employment models had lasting impacts on organizational employment pat-

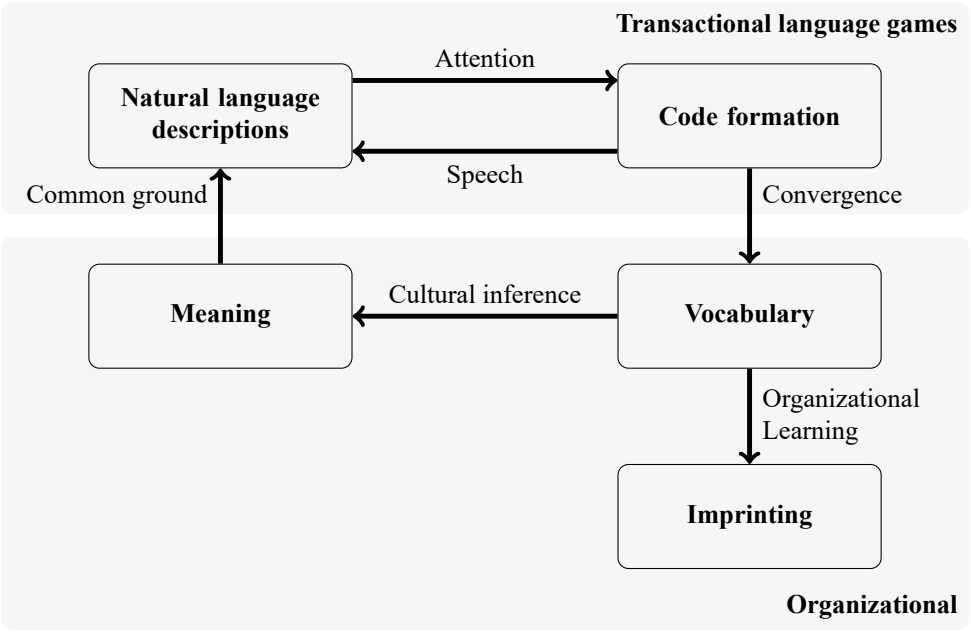


Figure 4.1. Process model for vocabulary formation and attention direction in the experiment.
The model illustrates two aspects of the experiment. The region labelled "Transactional" indicates things that occur when individuals engage with each other in a specific context with a specific goal in mind, and the region labelled "Organizational" indicates things that persist between transactions and, hence, are ontologically organizational.

terns and managerial intensity (Hannan, 2010). Phillips (2005) demonstrated that founders' blueprints influenced gender hierarchies in law firms, with founders from parent firms promoting women into partnership positions in progeny firms. Additionally, (Burton, Sørensen, and Beckman, 2002) and (Boeker, 1989) highlighted imprint transfer, wherein founders from prominent firms pursue innovation-based strategies, and initial choices impact organizational structure, respectively. A four-decade study by Dimov, de Holan, and Milanov (2012) showed that venture capital firms' initial decisions to invest early in high-technology industries strongly imprint subsequent investment decisions, particularly for firms with greater investing experience at the time of the first decision.

4.2.4. Model

This Section presents a model that explains how vocabularies are imprinted on organizations. The model shows how vocabularies, defined as a code schema, are formed through the repeated performance of a task that requires an actor with an object to convey to others in the same organization in which category that object sits. This could happen, for example, when an organization processes customer feedback, categorizes employee complaints (as harassment or amusing banter, for exam-

ple), or decides how to classify equipment failures. The group of people that perform this task is fixed in the short-term, but changes slowly over time.

Our model is illustrated in Figure 4.1. The model identifies activities that occur at a *transactional level*, through language games, and at a *organizational level*. Transactional level phenomena arise when individuals transact with each other: that is, when they speak, exchange things, and otherwise engage in specific situations in pursuit of their goals. Organizational phenomena are the organization's connective tissue: they extend beyond individual situations and so provide a common ground for the organization's transactions. These phenomena become imprinted in the organization and can outline any individual member.

Transactional language games

Transactional level activities occur when organizational members engage with one another to accomplish a specific task (Tomasello, 2009; Ocasio, 1997). Recall that each task requires one organizational member to convey to others in which category a particular object sits. The process by which that occurs is illustrated at the top left of the Figure. It occurs as individuals within the organization engage in *language games*. Language games bundle words and actions in commonly understood ways that allow people to act on the world together. They emerge as people use natural language to focus their attention upon task features that can be encoded and, in turn, use those codes as a basis for their natural language communications.

Language games According to Wittgenstein (1953), language acquires its meaning not because it comprises words with a commonly understood set of referents, but because groups of people know how to use language to get things done. Linguistic meaning is therefore both contextual and social. For example, Wittgenstein claimed that the phrase "five red apples" is meaningful in a grocer's shop because of the actions it triggers: specifically, it induces the the greengrocer to identify apples that are red and then to recite the sequence "one, two, three, four, five" while transferring the apples to a paper bag, before accepting payment (Wittgenstein, 1953, Para 1). In another context (for example in a computer store), the phrase would have a different effect and

so would have a different meaning. Wittgenstein's insights have generated a vast literature on the ways that communities create lexicons that help them to get things done in very specific contexts (Clark, 1998; Clark, 1996; Fussell and Krauss, 1992; Clark and Wilkes-Gibbs, 1986).

Our model envisages that language games develop in the following way. First, suppose that organizational members lack a standard way to assign an object to a category. In this case, they are forced to rely upon natural language. As members discuss the object under consideration, their language reflects the aspects of the object that they consider in order to assign it to a category. As indicated by the arrow from the "Natural language" node to the "Code formation" node, the result is that organizational members come to attend to the same aspects of their environment (Chun, Golomb, and Turk-Browne, 2011). As a result, organizational members start to coordinate on commonly understood codes. As indicated by the arrow labelled "speech," organizational codes are used in natural-language communication. In turn, that communication serves to reinforce and further to refine the organization's codes.

Code formation. The second stage in language game development occurs when the attention selection that occurs in natural language communication causes actors to pick out similar aspects of the object and to develop common codes to refer to them. Over time, the use of those codes starts to feel natural and so is reinforced.

The process that we have outlined so far could arise as any group of individuals work together to classify objects. We now turn to the features of the organization that allow the language games that inform object classification to outlast the tenure of specific group members.

Organizational-level imprinting

Organizational activities at the transactional level arise in specific instances of an object categorization task; organizational level features of the organization exist apart from specific transactions and so are persistent. One can therefore view organizational features of an organization as constitutive of its identity, because they survive even as the organization's personnel change and its assets turn over.

Our organizational concerns have their antecedents in problems that date back to the Ancient Greek philosopher Plutarch, who asked whether the ship sailed by Theseus, the founder of Athens, remained Theseus' ship a hundred years after his death, when all of its planks and sails had been replaced piecemeal (Scaltsas, 1980). Unlike Plutarch, we attempt to identify the organizational features of an organization: as indicated in the Figure, we claim that they relate to its vocabulary and to the meanings that it creates. Our model therefore underpins a linguistic and cultural understanding of organizational identity.

Vocabulary. The building block of the transcendent part of the organization is a vocabulary (for a related claim, see Cremer, Garicano, and Prat (2007)). As organizational members repeatedly transact, they start to generate common language games. In the specific context we have in mind, the organization's central task is to assign objects to categories and, consequently, the emergence of common language games bottoms out in an understanding that a given category is associated with a particular word (Nigam and Ocasio, 2010; Kennedy, 2008; Loewenstein and Gentner, 2005). The resultant set of (category, word) pairs is the organization's vocabulary (Loewenstein, Ocasio, and Jones, 2012).

The process by which transactional-level engagement generates an organizational vocabulary is indicated in Figure 4.1 by the arrow labelled "convergence." Convergence happens over repeated iterations of a given transaction, and it need not result in a completely stable vocabulary. Hence, as indicated in the Figure, a nascent vocabulary provides the social common ground for future transactions, which, in turn, serves to update the vocabulary. As the vocabulary emerges as an organizational feature of the organization, the organization can be conceptualised as a community of interpretation that understands a word to refer to a category whose members are themselves understood by the organization (Gentner, 2010; Waxman and Markow, 1995). For example, as discussed above, (coffee shop, McDonalds) might be part of an organization's vocabulary.

An established vocabulary becomes imprinted on its organization and it underpins shared organization-level meanings. Each of these is indicated by a node in Figure 4.1; we discuss them in turn.

Imprinting. Imprinting occurs when an organization acquires its own vocabulary that is passed down through successive generations of organizational members. When people in an organization

learn its vocabulary, they internalize a top-down schema that drives their response to categories and so affects their recollection of those categories. That is, as discussed in Section 4.2.1, a vocabulary furnishes people with a way of behaving when stimuli are processed in working memory. Organizational members respond to a category by bringing the associated word from their vocabulary to mind (Wickens, 1970). As a result, their recollection of the category is tilted towards the word's referent.

Meaning. When a vocabulary becomes imprinted upon an organization, successive generations of organizational members naturally learn to perform the cultural inference that assigns a community-specific meaning ("coffee shop") to a word ("McDonalds") (Loewenstein, Ocasio, and Jones, 2012; Colunga and Smith, 2005; Malt, Sloman, and Gennari, 2003).

Of course, organizational facts start to be created as soon as the first transaction occurs. The organization's set of organizational facts emerge from that time, and they continue to evolve over time. But, at every transaction after the first, they serve as a common ground of understanding and attentional perspective that serves as a backdrop to the performance of language games and transaction-level attention selection (Loewenstein, Ocasio, and Jones, 2012; Okhuysen and Bechky, 2009; Feldman and Pentland, 2003; Bechky, 2003; Cramton, 2001; Clark, 1998; Clark, 1996; Clark and Wilkes-Gibbs, 1986). This fact is indicated by the arrow labelled "common ground" in Figure 4.1. The fact that that common ground exists renders it very hard to change vocabularies and the associated attentional perspective and cultural meanings (Weber and Camerer, 2003).

We have now outlined the relationship between language, culture, and attention direction, and we have presented a simple model that relates the language games played in specific organizational contexts to the transcendent features of the organization that persist as its personnel and assets change over time.

4.2.5. Hypotheses

In this Section, we use our model of language games and vocabulary imprinting to develop empirical hypotheses.

As indicated on the vertical arrow on the right of Figure 4.1, our model predicts that language games within specific transactions cause convergence on a vocabulary that exists at the organizational level. In turn, as indicated by the vertical arrow on the left of the Figure, the vocabulary generates organization-level meanings that anchor future transactions and so cause further linguistic convergence. That means that, over time, vocabularies that outlast organizational members are imprinted on organizations. This argument is summarized in Hypothesis 4.1:

Hypothesis 4.1 (Convergence). *Repeated iterations of a transaction imprint vocabularies on an organization: that is, the organization converges on a vocabulary that outlasts its members.*

Precisely because vocabularies underpin meanings that provide a common ground for the language games that organizations use to perform transactions, the time that it takes to perform a transaction should drop as the vocabulary develops:

Hypothesis 4.2 (Performance). *Imprinting a vocabulary on an organization causes it to operate more efficiently: that is, as vocabularies converge, the speed with which organizations perform transactions diminishes.*

Our model explicitly incorporates "organizational" features. Recall that an organization's organizational features exist apart from specific transactions and so are persistent. They are also explicitly *organizational*. That is, by our account, an actor inside an organization recognises its organizational features when undertaking specific transactions; the same actor should therefore behave differently if the organizational features of his or her organization change:

Hypothesis 4.3 (Organizational identity). *Imprinting is fundamentally* organizational*: that is, when people move between organizations, their task performance in the new organization is more likely to be informed by that organization's vocabulary than by vocabulary of their old organization.*

We now turn to the relationship between the cognition of organizational members and the way that they experience the organization's language. We hypothesise that the way in which organizational members experience the world reflects, and in turn affects, the words that are used within an organization to describe a particular category.

Hypothesis 4.4 (Vocabulary imprint). *For any image, the likelihood that a vocabulary word predicts a given subject's cognition of an image is:*

- *Increasing in the number of words spoken*
- *Increasing in the amount of time spent*
- *Higher if the subject first encounters the word in Organization 1.*

We are also concerned with the long-run influence of transactional language. Recall from Figure 4.1 and the associated discussion that, once an organization starts to acquire an attentional perspective, that perspective informs its transactions, and those transactions further influence its attentional perspective. If this reasoning is correct, then founder language should have a long-run effect on attention direction:

Hypothesis 4.5 (Founder imprinting). *The language used by a founder predicts organizational members' cognition of the world after the founders leave the organization.*

4.3. Experiment

In this Section, we outline the empirical strategy that we use to test our hypotheses. We first explain the challenges that make it hard to perform a test in a live organization. Those challenges led us to an experimental approach, which we detail below.

4.3.1. Empirical challenges

Our model establishes a number of relationships. It suggests that natural language use causes organizational vocabularies to emerge; that vocabularies identify and create categories that structure group cognition and, thus, inform action; and that, once imprinted, vocabularies have long-lived effects on the way organizational members perceive the world. This Section explains why it is hard to test these relationships in live organizations and, hence, justifies our decision to use an experimental set-up.

The first challenge is a completely standard one in studies of culture: namely, that it is very hard to separate the effects of the phenomena under investigation from the shared understandings

and cultural assumptions that people acquire simply by virtue of working together, of sharing social engagements, and of living in close proximity. Any study of vocabulary formation and evolution in live organizations would have to explain why vocabularies and their meanings derived from the specific organizational factors that interest us.

The second challenge is that, as a practical matter, it is very hard to determine the structure of an organization's vocabulary and, hence, to study its evolution. The reason is that the researcher almost certainly cannot capture and record every relevant interaction between organizational members: many important conversations happen in non task-oriented contexts.

Third, vocabularies are only one of several things that could inform actors' choices: other social signals, or social conventions, might modulate the meaning of words, or simply override those words. As a result, a clean identification of the effects of a vocabulary is very hard.

Fourth, we are in general unable to follow an organization, its vocabularies, and the ways in which its members cognition is directed, from the inception of the organization. This makes it hard to substantiate claims about the origins of vocabularies. It also renders it very hard to study questions about the longevity of organizational features.

Fifth, as Figure 4.1 indicates, the relationship between categories, vocabularies, and organizational imprinting is complex and reflexive. Categories are caused by language games. Vocabularies rrstandings of categories and affect the ways in which language games evolve and are imprinted on organizational members after the founders have left. It would be far easier to study causation if some of this complexity were removed.

And, finally, the study of persistence can only occur in the presence of staff turnover. But staff turnover occurs for a variety of complex reasons that might themselves affect organizational features. Once again, in a first stab at this problem, it is reasonable to attempt to circumvent this complexity.

These features render it very hard to examine the phenomena with which we are concerned in a live organization. It is for this reason that we conduct our research in experimental organizations. Our experimental organizations are endowed with categories, working practices, and success measures. We therefore do not have to control for the emergence of these organizational features, and

variation in those features does not undermine cross-organizational comparisons. Staff leave our organizations automatically after a fixed period and are never fired for underperformance; furthermore, all people are promoted through our experimental organizations at exactly the same rate. And, finally, our experimental subjects have never met before they join their organizations, and they never interact outside those organizations.

4.3.2. Experimental design

We have now discussed the relationship between attention and organizational languages. We have presented a theoretical framework that relates, on the one hand, the language games that underpin the transactions that organizational members use to assign objects to categories to, on the other hand, the transcendent organizational artefacts of vocabulary, culture, and meaning. We have also argued that the most efficient way to study this relationship is experimentally. We now present an experimental set-up in which members assign objects to pre-defined categories and in which it is possible to track the evolution of language and meaning, and also to study attention direction.

The subjects in each experiment were assigned to one of two experimental organizations, which we referred to as the stand-alone and the cross-pollinated organizations. Each organization comprised three people: the director, and two staff members: the associate, and the trainee. It is convenient to assign the respective pronouns “she”, “he”, and “it” to the director, associate, and trainee. The organization performed several tasks. A single task comprised twelve picture-matching activities. The picture-matching activities deployed a fixed gallery of twelve pictures (Figure 4.B.7 in the Appendix). In each picture-matching activity, the director could see one of the gallery (Figure 4.B.5 in the Appendix) and was required orally to describe the picture to the staff members, who could see a common screen containing the entire gallery (Figure 4.B.6 in the Appendix).

We interpret each picture in the experiment as a category. Viewed in this way, the game requires the director to use words to help the associate and the trainee to identify which category she or he is looking at. Every picture was painted by the Dutch painter Pieter Bruegel the Elder. The paintings that we used are reproduced in the Appendix. Bruegel’s painting are famously complex: each of the paintings that our experimental subjects saw contained multiple identifiable features. For example,

when the subjects looked at *The Census at Bethlehem* (1566), they saw two people skinning a hog; three chickens pecking in the snow; three barrels on wheels, a child holding its mother's hand, tax collectors standing at a window, a man gathering firewood, a frozen lake, a hanging wreath, a red setting sun, two people hauling a barrel across a frozen lake, a line of people walking across the lake, children playing snowballs, people building a house, and so on and so on. When the picture is viewed as a category, that category contains all of these things. When we present the same picture to multiple organizations, we can compare the words that they assign to the picture in their vocabulary, as well as the ways in which the meanings that they assign to the picture vary across time.

The picture recognition game was timed. During the activity, the trainee could manipulate the staff screen but could not speak; the associate could speak but could not manipulate the screen. When the staff believed that they had identified the task picture described by the director, the trainee selected the image. If the image selected was correct, then the clock was stopped to yield an image time, measured in seconds, and the image on the director's screen changed. If the trainee submitted an incorrect picture, then a 10-second penalty was levied and the staff continued to try to identify the image. Any rule violation also resulted in a 10-second penalty.

When all twelve task images had been identified, the time taken to complete the task was revealed and each member of the organization was asked to estimate the completion time for the next task. Each member of the organization received a payment equal to £1 minus the sum of the image times divided by 180; the minimum payment for any task was £0. The baseline experiment comprised a learning round and six continuation rounds. The learning round occurred at the start of the experiment and comprised 10 experimental tasks; the continuation rounds each comprise 5 experimental tasks. Subjects were randomly assigned to experimental roles at the start of the learning round.

At the end of each round:

1. The organization's payoff for the round was revealed;
2. Each member of the organization evaluated the performance of the others on a Likert scale;
3. The trainee was promoted to associate;

4. The associate was promoted to the director's position;
5. The director of the standalone organization became the trainee in the cross-pollinated organization;
6. A new experimental subject entered the standalone organization as trainee;
7. The director of the cross-pollinated organization retired;
8. Every member of the organization predicted how many seconds the next task would take;
9. Each subject also received an updated version of the role illustration that highlighted their upcoming role (Figure 4.B.4 in the Appendix).
10. All subjects were required to click a button to indicate their readiness to start the next round.

When they exited the game, subjects participated in a memory task. They were shown each of the twelve images, in a random order, for 2 seconds. After they had been shown an image, they described the image and were paid 5p for each of the pictorial elements (e.g., tree, house, couple) and the details of those elements (e.g., brown wheel covered in snow; red house with thatched roof; couple sitting by stream; brunette woman wearing blue dress) that they recalled.

The experiment was originally designed to run in a physical lab environment but, due to Covid-19 restrictions, it was adapted to an online environment. The revised experimental design called for subjects to meet in a Zoom breakout room; the experiment was then conducted using a software environment developed in oTree. All subjects were sent a link to the Zoom meeting and the oTree platform; they were asked to read an information sheet (Figure 4.B.1 in the Appendix), to fill out a consent form (Figure 4.B.2 in the Appendix) and to read detailed instructions (Figure 4.B.3 in the Appendix) prior to being admitted to an organization or placed in a waiting room. Subjects were also given a diagram illustrating how the experiment worked (Figure 4.B.4 in the Appendix).

4.4. Results

Our experimental sessions were run from late-2021 to mid-2022. Each session comprised a single iteration of the two-organization experiment described above; on average, a session took 3.5 hours to run and generated a transcript that ran to 120 pages of text (approximately 30,000 words). We ran

20 sessions with 260 subjects through the Centre for Experimental Social Science at Nuffield College in the University of Oxford. The experimental hypotheses were pre-registered (AsPredicted registration number 32380). Two hypotheses were edited thereafter and included in the ethics application to CESS before the experimental sessions took place.

At the beginning of each session, participants were sent unique links to the experiment portal and Zoom via the CESS platform. They all consented to being recorded for the duration of the experiment. At the end of every session, the audio files were transcribed by the authors. The first draft was produced by an automatic transcription service (Otter.ai) and edited manually to verify accuracy. Each subjects' speech was identified by their unique alphanumeric identifier, their experiment, their experimental round, task, and image within the round. For the purposes of analysis, we defined an organization's task number to be an integer from 1 to 480 that orders every image that the organization identifies temporally. We performed standard pre-processing to produce unigrams and bigrams, which will be our main unit of analysis. [1]

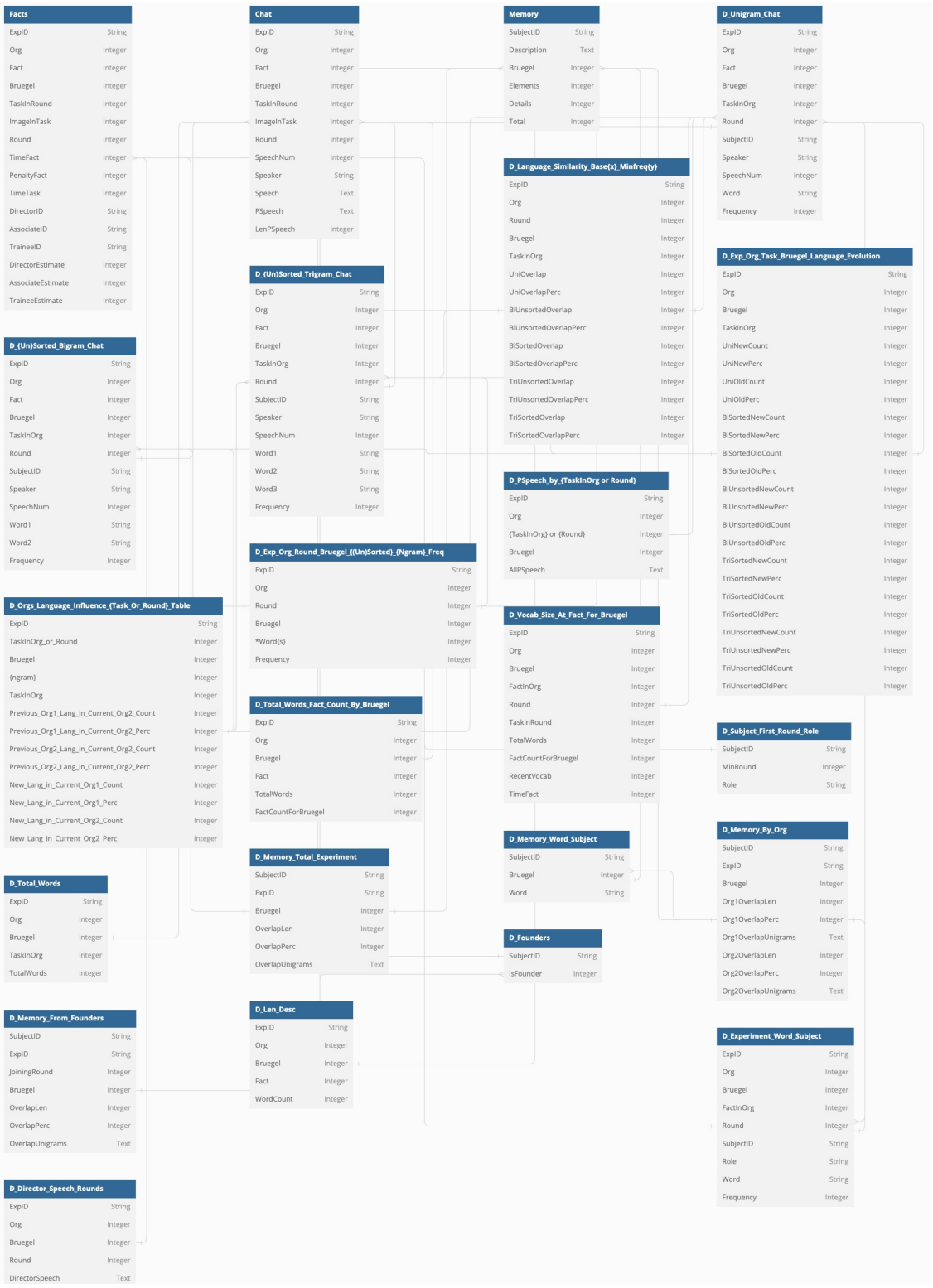
Moreover, because remuneration depended on the number of features and details recalled during the memory task completed by participants at the end of the experiment, the authors manually counted each memory task. A member of staff at CESS did the same. If there was more than a 2-point difference between the two scores, a third person would arbitrate. Note that the participant information was omitted from this stage to avoid bias. The transcripts, memory task scores, participant information, and performance information was subsequently processed and inserted into a MySql database for interrogation and analysis. Figure 4.2 shows all the tables in the database made of raw or processed data.

The images used for the experiment were pre-tested in MTurk. 200 users were shown 30 paintings by Pieter Bruegel the Elder and asked to rate the level of complexity of the image on a sliding scale from 0 to 100. We averaged all responses, ranked them, and chose to use the 12 images in the middle. This was necessary to ensure that the images used were of a comparable level of difficulty.

[1] An n-gram is a contiguous sequence of n tokens or words from a give sample of text or speech. The "n" in n-gram represents the number of items in the sequence. In this paper, our analysis uses unigrams and bigrams, which contain one and two words respectively.

Figure 4.2. MySQL Database Design

Figure 4.2 shows the database design in MySQL. Tables *Chat*, *Facts*, and *Memory* contain raw experimental data. All tables that start with ”D” are derived tables.



4.4.1. Convergence

Hypothesis 4.1 (Convergence). *Repeated iterations of a transaction imprint vocabularies on an organization: that is, the organization converges on a vocabulary that outlasts its members.*

Recall that an individual perceives a consensus when their primary expectation is that a group shares a similar code (Zou, Tam, Morris, Lee, Lau, and Chiu, 2009). Strong cultures therefore use highly similar and precise codes (Chatman, Caldwell, O'Reilly, and Doerr, 2014). We hypothesize that, as the task is performed repeatedly, participants converge on organizational codes that focus their *cognition* on specific features of the image and allow for faster image identification. We expect that both the number and percentage of old words (i.e. words used in previous tasks) will increase over time.

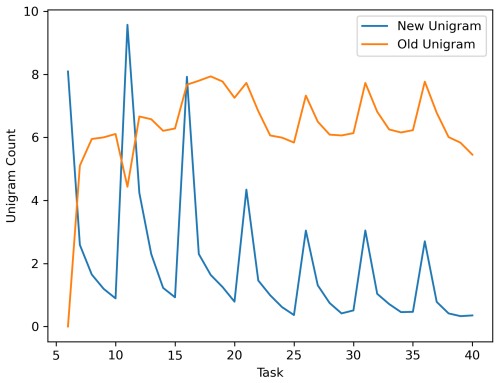
As anticipated, our experimental subjects rapidly developed organizational codes that supported their task performance. Each time an unigram U is used in organization $O \in \{1, 2\}$, we classify that usage as *new* if U has never been used in organization O; the unigram is classified as *old* if it has previously been used in organization O. In order to study the evolution of organizational languages, Figure 4.3 presents data on new and old unigram usage in Organizations 1 and 2.

In each of the sub-figures of Figure 4.3, the number of new unigrams in each organization spikes when a new member joins the organization: because the new member needs to learn the organization's vocabulary, experienced members of the organization use more words to explain that vocabulary to the new arrival. In Organization 1 (Figures 4.3a and 4.3b) and Organization 2 (Figures 4.3c and 4.3d), the average number of new unigrams used per round rapidly converges to just under 2 in rounds where a new member joins the group and to zero in other rounds. In tasks where new members join, the number of old unigrams used converges to about 7 in Organization 1 and to less than 6 in Organization 2; in tasks with no new members, the corresponding figures are about 6 and 5.

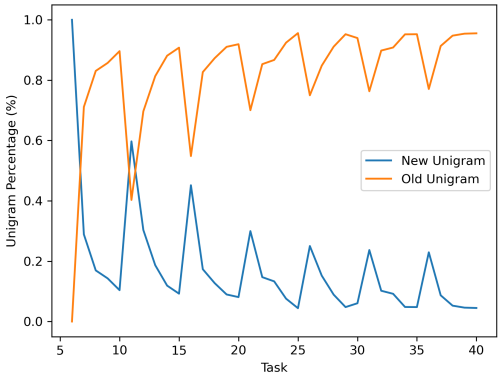
The results presented in Figure 4.3 suggest that language converges further in Organization 2 than Organization 1. Figure 4.4 presents graphical evidence of old unigram usage that supports

Figure 4.3. Unigram count and percentage by organization.
Figure 4.3 shows the number (count) and percentage of new and old unigrams by task. We extracted all words spoken by the Director in each experiment ID, organization, task, and image combination. We then identified whether each word was novel to an organization's vocabulary or whether it had been previously used (i.e. used in a previous task). We calculate the number of new words and old words and the percentage of new words and old words in each experiment ID, organization, task, and image combinations and then produce an average. We exclude the first 5 tasks.

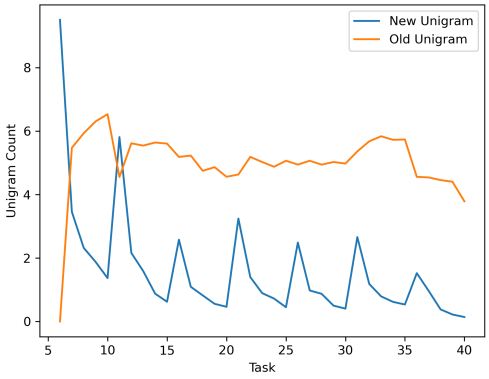
(a) Organization 1: New vs. Old Average Count



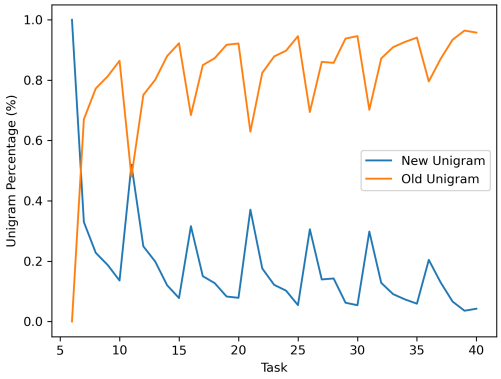
(b) Organization 1: New vs. Old Average Percentage



(c) Organization 2: New vs. Old Average Count



(d) Organization 2: New vs. Old Average Percentage

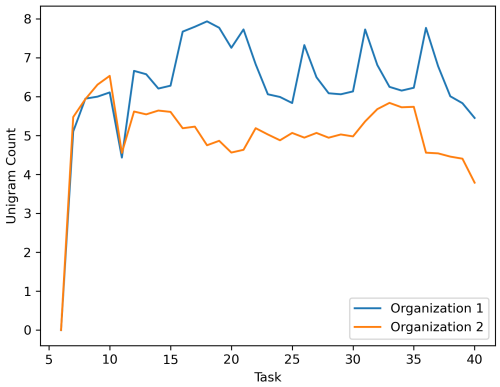


this conclusion. The number of old unigrams drops further in Organization 2 than Organization 1; the percentage old unigram usage is about the same in both organizations, because the size of Organization 1's language is larger than Organization 2's. Further evidence that Organization 2's language converges further than Organization 1's is presented in Figure 4.5, which shows that the total number of unigrams used in Organization 2 is lower than it is in Organization 1 after the 10th round of the experiment. As shown in Figures 4.A.1 and 4.A.2 in the Appendix, we conduct the same analyses for bigrams and the same pattern holds.

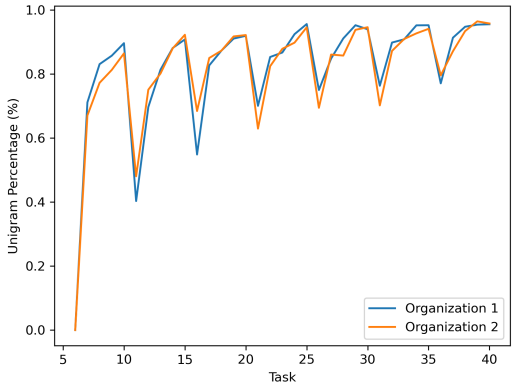
Table 4.A.1 in the Appendix presents the results of a difference-in-means test for the difference

Figure 4.4. Old Unigram Count and Percentage.
Figure 4.4 shows the count and percentage of old unigrams by organization. As per Figure 1, we extracted all words spoken by the Director in each experiment ID, organization, task, and image combination. We then identified whether each word was novel to an organization's vocabulary or whether it had been previously used (i.e. used in a previous task). We calculate the number (count) and percentage of old words in each experiment ID, organization, task, and image combinations and then produce an average. We exclude the first 5 tasks.

(a) Old Unigram Count



(b) Old Unigram Percentage



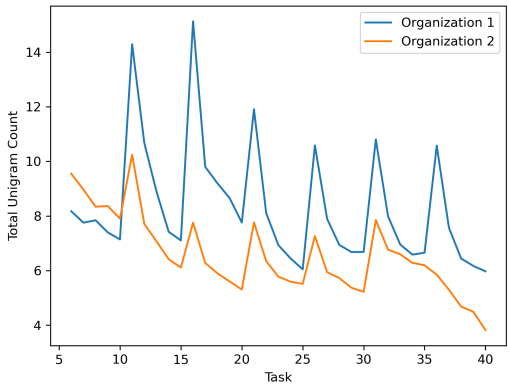
in the number of old unigrams in Organizations 1 and 2. In line with the graphical evidence of Figure 4.3, the number of old unigrams in Organization 2 is very significantly lower than the number in Organization 1 for almost every task after Task 12 (i.e. by the beginning of Round 2) at a 95-99% significance level.

We hypothesize that language converges more in Organization 2 than Organization 1 because, as the experiment progresses and members that were originally in Organization 1 join Organization 2, they can use the vocabularies developed in both organizations. Note, however, that while the count average total count of old unigrams in Organization 1 is higher than in Organization 2, the percentage of old unigrams is almost identical. This mechanically means that Organization 1 uses more total words than Organization 2, as is shown in Figure 3. Again, we ran a difference-in-means test and this difference is statistically significant across most tasks (see Table 4.A.1).

4.4.2. Performance

Hypothesis 4.2 (Performance). *Imprinting a vocabulary on an organization causes it to operate more efficiently: that is, as vocabularies converge, the speed with which organizations perform*

Figure 4.5. Total Unigram Count for Organization 1 and Organization 2.
Figure 4.5 shows the total average count of unigrams for Organization 1 and Organization 2. We extracted all words spoken by the Director in each experiment ID, organization, task, and image combination. We then identified whether each word was novel to an organization's vocabulary or whether it had been previously used. We total unigram count is the addition of new unigram count and old unigram count. Finally, we average total unigram count for across each organization and task. We exclude the first 5 tasks.



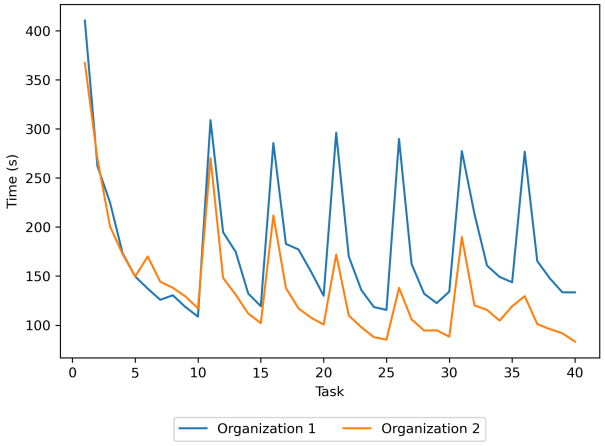
transactions diminishes.

The software developed for the experiment automatically recorded the time to identify each Image I in organization $O \in \{1,2\}$. Recall that each is task τ is composed of 12 image identifications. We write $t_O(\tau, I, e)$ for the time taken in organization O in experiment e during task τ to identify image $I \in \{1,2,\ldots,12\}$. Then the total task time for task τ in organization O in experiment e is defined to be $t_O(\tau, e) = \sum_{I=0}^{12} t_O(\tau, I, e)$. The *average task time* $t_O(\tau)$ for task τ in organization O is the arithmetic average of $t_O(\tau, e)$ across all experiments e. The *average round time* $t_O(\rho)$ for round ρ in organization O is the arithmetic average of $t_O(\tau)$ across all of the tasks τ that comprise round ρ.

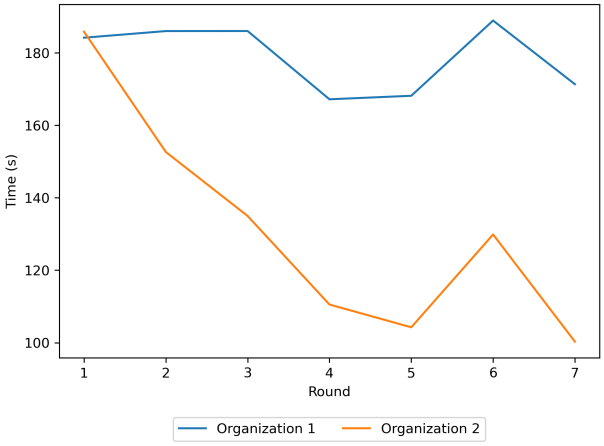
Hypothesis 4.2 claims that the speed at which organizations perform transactions diminishes as their vocabularies converge. Each of the twelve images $I \in \{1,2,\ldots,12\}$ used in the experiment appears once in every task τ in every round ρ of the experiment e. Figure 4.6 shows the average task time $\tau_O(\tau)$ and the average round time $t_O(\rho)$ that our experimental organizations take to identify each image. Each graph has one line that shows the average time taken to complete a task by organization $O \in \{1,2\}$. The Figure provides graphical evidence in support of the hypothesis: performance improves with experience, i.e. $t_O(\tau)$ falls over time. The same pattern applies if the sample is split by image I (see Figure 4.A.3 in the Appendix).

Figure 4.6. Average time per task.
Figure 4.6a shows the average time taken to complete a task. As written in the text, we use $t_O(\tau, I, e)$ for the time taken in organization O in experiment e during task τ to identify image $I \in \{1, 2, \ldots, 12\}$. Then the total task time for task τ in organization O in experiment e is defined to be $t_O(\tau, e) = \sum_{I=0}^{12} t_O(\tau, I, e)$. The *average task time* $t_O(\tau)$ for task τ in organization O is the arithmetic average of $t_O(\tau, e)$ across all experiments e. The *average round time* $t_O(\rho)$ for round ρ in organization O is the arithmetic average of $t_O(\tau)$ across all of the tasks τ that comprise round ρ. At the end of each round, there is a staff turnover: the director leaves the organization, a new trainee enters the organization, and the previous trainee and associate are promoted.

(a) Old Unigram Count



(b) Old Unigram Percentage



As above, the speed with which an organization identifies an image spikes up at the beginning of every experimental round ρ. We believe that the reason for this is that the organization's personnel changes at the end of every round and this turnover temporarily impairs performance: the director leaves, the trainee and the associate are promoted, and a new trainee arrives. The new organization has to learn how to operate, and the new trainee has to acquire the organizational language that the director and the associate are using. Even accounting for this performance impairment, the average task time by round $t_O(\rho)$ falls across rounds for Organization 2 but remains relatively stable for Organization 1. After the initial learning phase ($\rho = 1$), Organization 2 consistently outperforms Organization 1. As shown in Table 4.A.2 in the Appendix, we ran a difference-in-means test and this difference is statistically significant; this applies to the average task time by round and average task time by task.

We use a regression analysis to quantify the factors that affect task performance. Across the entire universe of tasks performed by organizations in our dataset, we regress the time $t_O(\tau, B, e)$ that organization $O \in \{1, 2\}$ takes to identify image I in a given experiment e against the O, I, τ

Table 4.1. The relationship between performance, progress through the experiment, and vocabulary size.

This table reports OLS estimates of the dynamics of TimeFact on RecentVocabulary, Organization, Task, and Image based on the following regression:

$$Time_Fact = \beta_0 + \beta_1 Recent_Vocabulary + \beta_2 \text{Organization} + \beta_3 Task + \beta_4 Image + \epsilon$$

As described in the text, *Time_Fact* is the time in seconds (s) taken to identify each image *I*. *Recent_Vocabulary* is the total number of unique words used by a given organization *O* in the previous 5 tasks τ. *Task* is the $\tau \in \{1, 2, \ldots, 40\}$ is used in organization $O \in \{1, 2\}$. Finally, *Image* is the specific *I*.

Model:	OLS
Dependent Variable:	*Time_Fact*
	(1)
Intercept	16.7106***
	(24.977)
Recent_Vocabulary	0.3037***
	(16.485)
Organization	-3.4617***
	(-11.958)
Task	-0.0792***
	(-5.035)
Image	-0.0141
	(-0.339)
Fit statistics	
R-squared:	0.030
Adj. R-squared:	0.030
F-statistic:	139.6
No. Observations:	17965

and *RecentVocab*, which we define to be the total number of unique words used by organization *O* to identify Image *I* in the previous 5 tasks; we start our regression at task $\tau = 6$, so *RecentVocab* is guaranteed to be well-defined. The regression equation appears in the caption to Table 4.1, where we report our results.

As anticipated, the time taken to identify a given image is very significantly decreasing in the task number τ. Organization 2 performs significantly better than Organization 1 across our entire dataset, and the time taken to perform a task is increasing in the *RecentVocab* for the relevant organization. These results are consistent with Hypothesis 4.2.

The model of Section 4.2.4 does not generate predictions about the relative rate of convergence for standalone and cross-pollinated organizations; the experimental work of Weber and Camerer (2003) identifies a long-lived detrimental effect of cross-pollination upon organizational performance. Figure 4.6 compares the performance of the standalone organizations (Organization 1 in

the Figure) and the cross-pollinated organizations (Organization 2 in the Figure) in our experimental set-up. For each round and each task in the experiment, the Figure shows the average time that each organization takes to identify all of the images. As in Figure 4.5, both organizations get better at identifying images as they gain experience and, also as in Figure 4.5, each organization suffers a performance drop when its staff turns over at the end of an experimental round. Strikingly, however, after the initial rounds of learning, the cross-pollinated Organization 2 consistently outperforms Organization 1. This result is not consistent with Weber and Camerer (2003). We hypothesise that the difference arises because our organizations have an imposed structure that allows new trainees to learn their organization's language before they attempt to use it. As a result, the type of confusion that arises when Weber and Camerer's subjects are forced simultaneously to use different learned languages to perform a task cannot occur in our experiment. Instead, a subject who transfers from Organization 1 into the Organization 2 gets to learn the cross-pollinated language and then to decide whether, as associate, to attempt to introduce new terms into that language. As a result, the cross-pollinated organization can take advantage of the best parts of both organization's languages.

4.4.3. Organizational identity.

We now turn to Hypothesis 4.3:

Hypothesis 4.3 (Organizational identity). *Imprinting is fundamentally* organizational*: that is, when people move between organizations, their task performance in the new organization is more likely to be informed by that organization's vocabulary than by vocabulary of their old organization.*

Organization 1 is a standalone organization in our experimental set-up and Organization 2 is cross-pollinated with staff from Organization 1 precisely in order that this hypothesis can be tested. We identify two ways to compare the use of organizational and transactional features of the organization. Both approaches rely upon an analysis of the language that people use in each organization: we study the language used in a given round in Organization 2, and we determine what fraction of that language reflects the transactional experience that the members of the Organization 2 have in

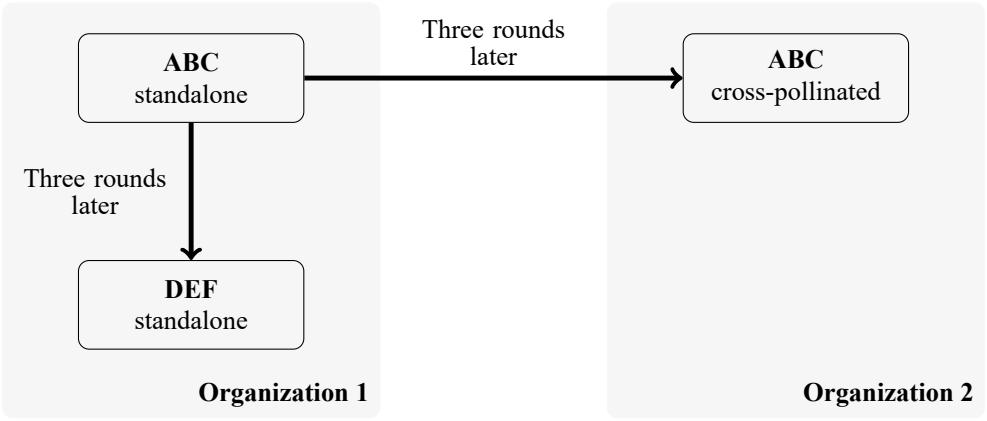


Figure 4.7. First test for organizational identity.
Our set-up allows us to place three people in the same roles to perform the same tasks in two different organizations. If actors A, B, and C occupy the respective roles of director, associate, and trainee in Organization 1 then, three rounds later, they will occupy the same roles in Organization 2, while new actors D, E, and F will occupy those roles in Organization 1. They could use the vocabulary of Organization 1 or Organization 2 to perform tasks. If organizational identity affects their choice, then the language used by A, B, and C in Organization 1 will be closer to the language used by D, E, and F than it is to the language that A, B, and C use in Organization 2.

Organization12, and what fraction is derived from the Organization 2.

The rationale for our first test is illustrated in Figure 4.7. It shows the evolution of the experiment from a round in which agents A, B, and C work together in Organization 1. Three rounds later, three different people comprise Organization 1: call them D, E, and F. At the same time, A, B, and C will meet again in Organization 2. If Hypothesis 4.3 is correct, then, because organizational features are more important than transactional experience, the language that DEF use in Organization 1 should be closer to the language that ABC used Organization 1 than is the language ABC use in Organization 2.

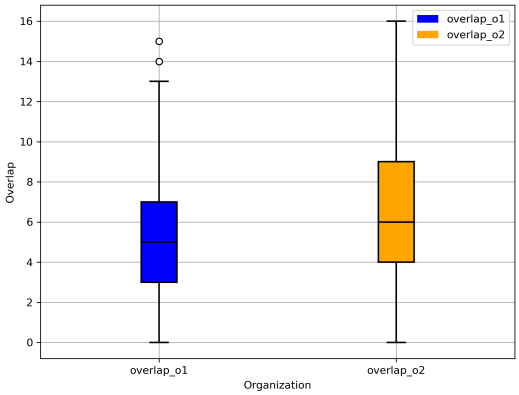
We establish whether the languages in the horizontal ABC organizations of Figure 4.7 are closer than the vertically arranges ABC and DEF organizations as follows.

1. For each experiment, we identify every base round for which it is possible to draw Figure 4.8. For example, if the base round is 1, then Organization 1 is comprised of ABC in round 1 ($\rho = 1$) and it is comprised of DEF in round 4 ($\rho = 4$); Organization 2 in round 4 ($\rho = 4$) is comprised of ABC.

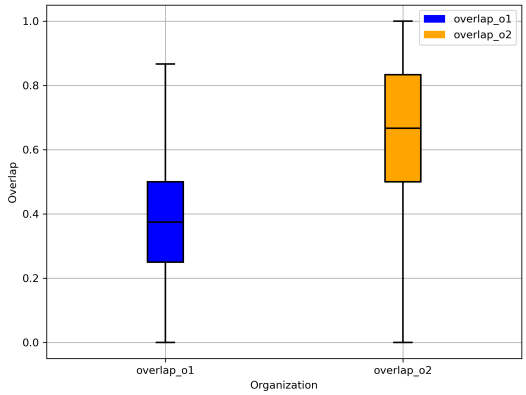
2. For each base round in each organization, we compute the list of words used by ABC_{O1}, ABC_{O2} and DEF_{O1}. We then compute the overlap between the words used by ABC_{O1} and ABC_{O2} and between the words used by ABC_{O1} and DEF_{O1}. We compute the number of words in both

Figure 4.8. Overlapping words in Organization 1 and Organization 2 three rounds after all base rounds.
Figure 4.8 shows the overlapping unigram count (a) and percentage (b) in Organization 1 and Organization 2 three rounds after all base rounds (namely rounds 4, 5, 6 and 7). Consider the meeting of three people in Organization 2 who have already met in Organization 1. For every such combination, we calculate the total unigram count and percentage that they learned in Organization 1 versus Organization 2 and average this across all Images and Experiment IDs. These results are displayed it box-and-whiskers plots. The top and bottom of the whiskers represent the maximum and minimum, respectively; the top of the box represents the upper quartile, the bottom of the box represents the lower quartile, and the line in the middle represents the median.

(a) Unigram count



(b) Unigram percentage



overlaps and also the size of the overlaps as a percentage of the number of words used in ABC_{O2} and DEF_{O1}.

3. We average the size and percentage of overlaps in the cross-pollinated and standalone organizations across all experiments and all valid base rounds.

The results of our calculations are reported as box and whisker plots in Figure 4.8. The overlap_o1 data is between ABC_{O1} and DEF_{O1}; the overlap_o2 data is between ABC_{O2} and ABC_{O1}. It is clear from the Figure that the latter overlap is significantly higher than the former and, hence, we reject Hypothesis 4.3.

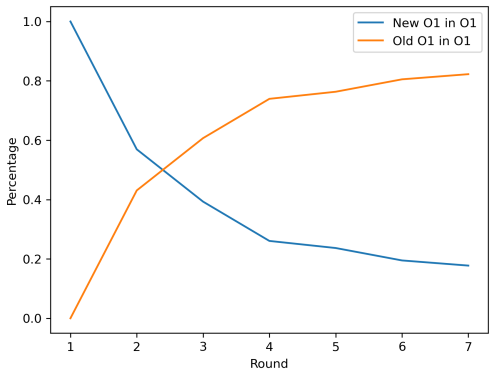
We have now described our first attack on Hypothesis 4.3, which examined whether the language that three people – A, B, and C – use in Organization 1 should be thought of as attached to Organization 1 or to those three people. We tested this by studying the difference between the language as it evolved over three rounds in Organization 1 and the language that A, B, and C use three rounds later in Organization 2. We now discuss our second approach to the Hypothesis.

Rather than studying the evolution of language from a starting round in Organization 1, our

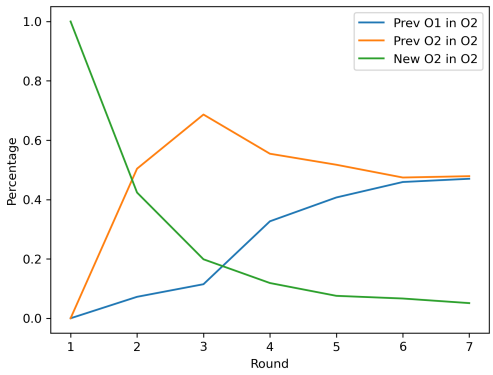
Figure 4.9. Percentage contribution to language in Organization 1 and Organization 2 by round and task.

Figure 4.9 shows the percentage of new and old unigrams by task. We extracted all words spoken by the Director in each e, O, and I combination by τ and R. We then identified whether each word was new (i.e. not been spoken in previous τ and R) or old (i.e. has been spoken in a previous τ and R) to an organization's vocabulary. We exclude the first 5 tasks and produce an average of across each combination described above.

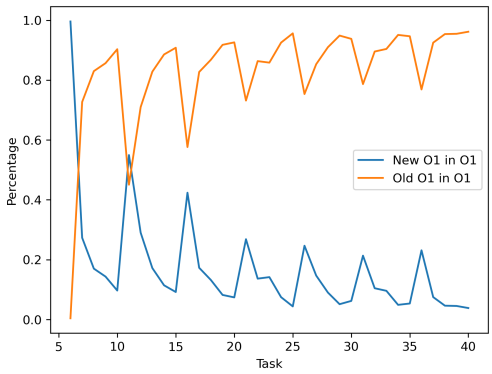
(a) Organization 1 by round



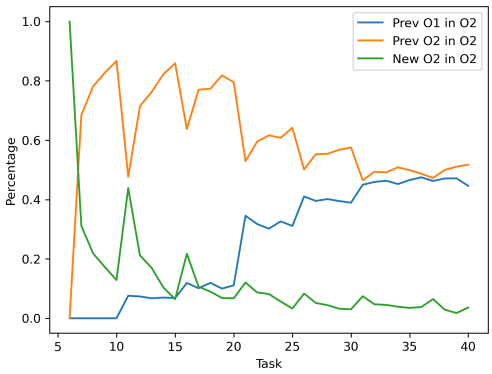
(b) Organization 2 by round



(c) Organization 1 by task



(d) Organization 2 by task



second approach to Hypothesis 4.3 looks backward from a given task or round and studies the source of the language used in that task or round. We perform two sets of analysis by both task and round.

In the following explanation, the counter c refers either to tasks or to rounds:

1. We perform the following analysis for each experimental organization and for counter c equal to the task $\tau = 6, 7, \ldots, 40$ or the round $\rho = 1, 2, \ldots, 7$.
2. We record all of the words used in Organization 1 for a given c; we classify them as new to Organization 1 for this c, or old.
3. We record all of the words used in Organization 2 for this c. We classify those words as (i)

completely new in Organization 2; (ii) previously used in Organization 2; previously used in Organization 2 but not Organization 2.

Our procedure generates two Organization 1 variables. (i) *New O1 in O1*: percentage of words used in count *c* in Organization 1 that are new to Organization 1; (ii) *Old O1 in O1*: percentage of words used in count *c* in Organization 1 that that were previously used in Organization 1. These figures are averaged across all experiments and are plotted for *c* as round and task, respectively, in Figure 4.9a and Figure 4.9c. As illustrated in the Figures, over the course of the experiment, the percentage of new words used in Organization 1 drops steadily, and the percentage of old words levels out. This is consistent with convergence at the organizational level but, of course, reveals nothing about cross-organizational identity comparisons.

We also generate three Organization 2 variables. (i) *Prev O1 in O2*: percentage of words used in count *c* in Organization 2 that originated in Organization 1; (ii) *Prev O2 in O2*: percentage of words used in count *c* in Organization 2 that originated in Organization 2; (iii) *New O2 in O2*: percentage of words used in count *c* in Organization 2 that are new to Organization 2. These variables are plotted for *c* as round ρ and task τ, respectively, in Figure 4.9b and Figure 4.9d. As for the Organization 1 variables, the percentage of new words diminishes steadily for both types of *c* over the course of the experiment, and the percentage of old words flattens out. Strikingly, however, the percentage of words used in Organization 1 that originate in Organization 1 and that originate in Organization 2 are identical at the end of the experiment. That is, while Organization 2's language owes more to Organization 2's history at the start of the experiment, that influence slowly diminishes as Organization 2's unique identity is slowly attenuated by the influence of Organization 1.

Finally, using our second approach to Hypothesis 4.3 which looks backward from a given task or round, we study where the Director of Organization 2 learned the language s/he used to describe the images. Consider the meeting of three people in Organization 2 – A, B, and C – who have already met in Organization 1. We are interested in how many of the words that the Director of Organization 2 speaks were words that s/he learned in Organization 2 versus learned in Organization 1. There are three definitions of *Learned O1*. It could mean that (i) the Director spoke the word in

Organization 1, (ii) the Director heard in Organization 1 as Trainee or Associate, or (iii) the Director either spoke or heard the word in Organization 1. Mechanically, there is only one definition for *Learned O2*: the Director heard the word in Organization 2 as Trainee or Associate.

We establish whether the language used by a Director of Organization 2 was learned in Organization 1 or Organization 2 as follows:

1. For each subject s that fulfils the role of Director in Organization 2 in each experimental session e:
2. We generate the set $W(s,I)$ of all words that s uses when serving as director of Organization 2 to describe each of the images $I \in \{1,2,\ldots,12\}$.
3. We define the set $L^2(s,I)$ of words describing image I that s learns in Organization 2 to be all of the words that s hears as an associate or trainee in Organization 2.
4. We define the set $L^1(s,I)$ of words describing image I that s learns in Organization 1 in three ways:
   (a) $D^1(s,I)$: the set of words that s speaks as Director of Organization 1 when describing image I.
   (b) $H^1(s,I)$: the set of words that s hears as Trainee or Associate in Organization 1 when describing image I.
   (c) $DH^1(s,I) = D^1(s,I) \cup H^1(s,I)$: the set of all words that s speaks or hears in Organization 1 when describing image I.
5. The elements of $L^1(s,I) \cap L^2(s,I)$ could have been learned in either organization; we exclude them from our sets of learned words.
6. For each definition of $L^1(s,I)$, we determine the size of the sets $W(s,I) \cap L^1(s,I)$ and $W(s,I) \cap L^2(s,I)$ of words used by s to describe image I that were learned in Organizations 1 and 2, respectively.
7. Finally, we average our results across every Organization 2 in every experiment.

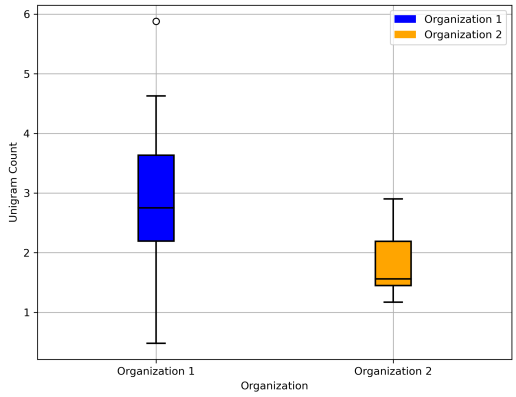
The results of the calculations above are reported as box and whisker plots in Figure 4.10. It is undeniable from the Figure that Directors in Organization 2 use more words both spoken in

Organization 1 and heard in Organization 1 than heard in Organization 2. Words spoken dominate over words heard. Note that the number of words stemming from Organization 2 is smaller in the latter case than in the former; this is a mechanical difference because the overlap in words heard in Organization 1 and heard in Organization 2 is bigger and these words are excluded, as outlined above.

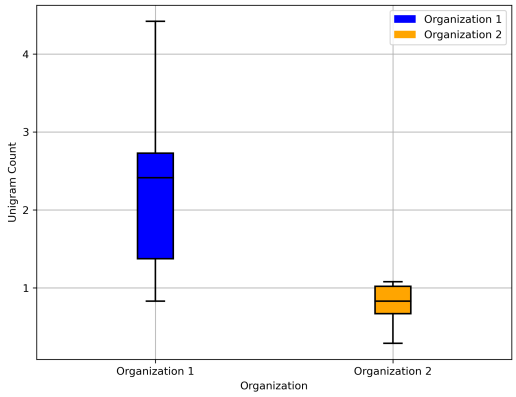
Figure 4.10. Words learned in Organization 1 and Organization 2

Figure 4.10 analyses how many of the words that the Director of Organization 2 speaks were words that s/he learned in Organization 2 versus learned in Organization 1. There are three definitions of *learned O1*. It could mean that (i) the Director spoke the word in Organization 1, (ii) the Director heard in Organization 1 as Trainee or Associate, or (iii) the Director either spoke or heard the word in Organization 1. Mechanically, there is only one definition for *learned O2*: the Director heard the word in Organization 2 as Trainee or Associate. We exclude all overlapping words that could have been learned in both Organization 1 and Organization 2. Comparisons between *learned O2* and each of the "learned O1" definitions are displayed it box-and-whiskers plots. The top and bottom of the whiskers represent the maximum and minimum, respectively; the top of the box represents the upper quartile, the bottom of the box represents the lower quartile, and the line in the middle represents the median.

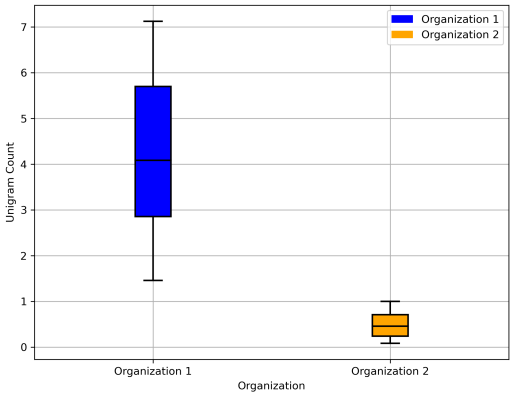
(a) Unigrams spoken in O1



(b) Unigrams heard in O1



(c) Unigrams heard and spoken in O1



Note that the results of the two tests are not incompatible. Rather, they offer an important nuance. The first test only compares an arithmetic average of language originating in Organization 1 and Organization across all the rounds, whereas the second test uses average by round. On average, *Prev O2 in O2* language dominates over *Prev O1 in O2* language. However, Figure 4.9 shows that the use of language originating in Organization 1 and Organization 2 coverges over time. By the final rounds, the difference is statistically insignificant. As more participants move from Organization 1 to Organization 2, they bring over more language imprinted during their participation in Organization 1.

4.4.4. Vocabulary imprint

Hypothesis 4.4 (Vocabulary imprint). *For any image, the likelihood that a vocabulary word predicts a given subject's cognition of an image is:*

- *Increasing in the number of words spoken*
- *Increasing in the amount of time spent*
- *Higher if the subject first encounters the word in Organization 1.*

Hypothesis 4.4 relates the cognition of organizational members and the way that they experience the organization's vocabulary. Recall from Section 4.2.1 that, because of the limited cognitive capacity of individuals, codes direct their cognition to specific environmental cues over others. When the external stimuli is withdrawn, we hypothesize that the likelihood that the participants would remember the experimental vocabulary is increasing in the number of words spoken because crisper vocabularies are a sign of strong cultural consensus (Kocak and Puranam, 2023). We also hypothesize that the likelihood of recall is increasing in the amount of time spent exposed to the image because increased exposure allows for greater sedimentation of the link between the external stimuli and the organizational vocabulary. Finally, we hypothesize that the likelihood of recall is higher if the subject first encounters the word in Organization 1 because the imprint happened earlier in the participants career, before they had any cognitive associations between the image and codes used to identify it. Individual cognition in our experiment is measured through a memory task that identifies the things that each subject recalls when shown a picture for two seconds after the game is been played. We contend that it is extremely hard to attend to any sensory stimuli during the two second presentation of the picture and, hence, the subject's cognition is directed not by the content of the picture as she or he perceives it during the memory task, but by the organizational-level schema that she or he absorbed while playing the game.

We compare the vocabulary used during the experiment with the vocabulary used during the memory task as follows:

1. For each subject s in each experimental session e:

2. We generate s's set $W(s, I)$ of *Experimental Words* for image I, comprising every word that s hears or speaks when image I is being described.
3. We generate sets $W(s, I, O1)$ and $W(s, I, O2)$ of words that s hears or speaks when image I is being described in Organizations 1 and 2 respectively. $W(s, I, O1)$ and $W(s, I, O2)$ can have a non-empty overlap.
4. For every experimental word $w \in W(s, I)$:
   (a) We define the *frequency* $f(w, s, I)$ of w for subject s and image I to be the number of times that s is exposed to w during a description of image I.
   (b) We record the *mode* $m(w)$ of s's encounter with w while image I is being discussed: $w(m, s, I)$ could be p (spoken), h (heard), or both.
5. For each image I, we calculate $t(s, I)$ to be the time that s spends on I and $T(s)$ to the time that s spends in the experiment.
6. For each image I, we record the set $M(s, I)$ of *Memory Words* that s uses when describing I during the memory task.
7. We calculate the size *unigram count* of the set $W(s, I) \cap M(s, I)$, which is the number of memory words that s encounters when I is being described; we also calculate the *Organization 1 unigram count*, $W(s, I, O1) \cap M(s, I)$, and the *Organization 2 unigram count*, $W(s, I, O2) \cap M(s, I)$.
8. We calculate size of each of the above overlaps as a percentage of *memory words*.

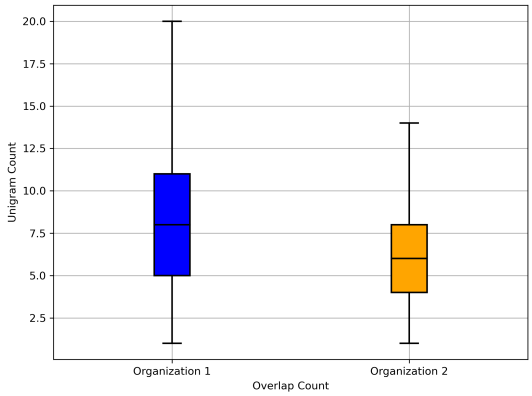
We start by looking at the impact of the Organization 1 and Organization 2 language encountered by a subject s ($W(s, I, O1)$ and $W(s, I, O2)$) on their cognition. Because subjects are incentivised during the memory stage of the experiment, *Memory Words* is a reliable proxy for their post-experiment recall. Hence, we use *Memory Words*, $M(s, I)$ as a measurement of cognition. We hypothesize that our three measures of *Experimental Words* direct the attention of subjects to certain aspects of the images I over others. Figure 4.11 shows that subjects recall more words from Organization 1 than Organization 2.

Table 4.2 uses difference-in-means tests to establish the statistical significance of the results illustrated in Figure 4.11. For each subject s, we compute s's *unigram count* of recalled words from Organizations 1 and 2 (see the above definition) and we determine these counts as a percentage of

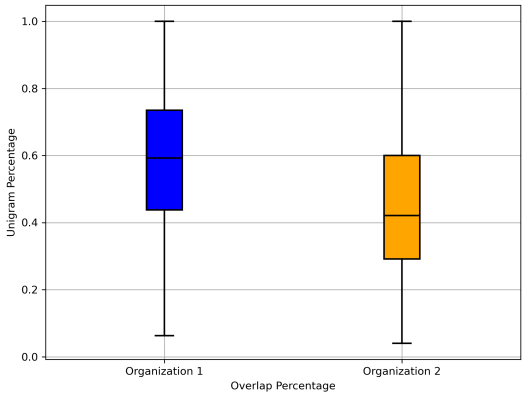
total words recalled. Those figures are averaged across all subjects; the Table reports difference-in-means tests for both measures for every Image.

Figure 4.11. Overlap of experimental words with memory words.
Figure 4.11 shows a box-and-whiskers plot of (a) the overlapping count and (b) the overlapping percentage of *Experimental Words* by *O* with *Memory Words*. For each *s*, we first extracted the *Experimental Words* per *O* and the *Memory Words* and then calculated the size of each overlap per *I* and the percentage this overlap represented of the *Memory Words* set. These results are displayed it box-and-whiskers plots. The top and bottom of the whiskers represent the maximum and minimum, respectively; the top of the box represents the upper quartile, the bottom of the box represents the lower quartile, and the line in the middle represents the median.

(a) Unigram count



(b) Unigram percentage



In order to analyze the likelihood that an experimental word is recalled during the memory task, we run a logistic regression of the probability of recalling a word during the memory task on the words spoken or heard by participants during the experiment.

We now investigate the likelihood that a word encountered by a subject during the experiment is recalled during the memory task. To accomplish this, we use a logistic regression to estimate the probability that a word w in subject s's set of experimental words $W(s, I)$ is recalled. Our independent variables are listed in the caption of Table 4.3, where we present our regression specification: the independent variables measure the total number of words spoken and heard by s during the experiment and for the image I, as well as the frequency with which the word w is spoken and heard by s.

The results of our regression are presented in Table 4.3. The probability of recall is decreasing in the number of words spoken or heard. The effect of spoken words is stronger than the effect of

Table 4.2. Difference-in-means of *Experimental Words* by *O* and *Memory Words*.
In Table 4.2, we calculate the difference-in-means of *Experimental words* by *O* with *Memory Words*. For each *s* and *I*, we first extracted the *Experimental Words* per *O* and the *Memory Words* and then calculated the size each overlap per *I* and the percentage this overlap represented of the *Memory Words* set. We subsequently average these results by *O* and *I* across all *s*. To calculate the difference-in-means, we subtract the $O = 2$ mean from the $O = 1$ mean, $\bar{O}_1 - \bar{O}_2$.

Image	Difference-in-Means	
	Unigram Count	Unigram Percentage
1	1.812***	0.117***
	(4.462)	(4.613)
2	2.315***	0.149***
	(6.011)	(6.243)
3	1.756***	0.120***
	(5.581)	(5.500)
4	2.307***	0.154***
	(5.567)	(6.536)
5	2.156***	0.127***
	(6.182)	(6.792)
6	2.573***	0.124***
	(5.762)	(4.733)
7	2.077***	0.151***
	(6.356)	(6.232)
8	1.846***	0.126***
	(5.359)	(5.587)
9	2.774***	0.168***
	(6.740)	(7.787)
10	3.070***	0.172***
	(7.565)	(8.519)
11	2.784***	0.186***
	(7.536)	(8.095)
12	2.588***	0.133***
	(6.365)	(6.406)
Overall	2.332***	0.144***
	(21.098)	(21.870)

heard words. The probability that a word is recall is higher if that word is spoken or heard more often. Again, spoken frequency has a higher impact than heard frequency.

We also run two further OLS regressions, presented in Table 4.4. The first model regresses recalled words *spoken* during the experiment, on total words spoken or heard during the experiment. The second model does the same thing with *unspoken* words, i.e. words that were never used during the experiment. Unspoken words are supposed to capture the elements of the pictures that the participants could have noticed outside the constraints of the organizational vocabulary. We find that spoken words are more likely to be recalled if more words are spoken and heard. Conversely, unspoken words are more likely to be recalled if less words are spoken and heard. The basic

intuition behind both these results is that if more language directs a participants cognition, then s/he has less bandwidth to explore the image outside the bounds of the organization's vocabulary. If less language directs a participants attentional perspective, than the opposite holds true. Taken together, these results provide support in favour of Hypothesis 4.4.

Table 4.3. Regression of the probability of recalling a word on words spoken or heard.
This table reports logit estimates of the relationship between the probability of recalling a word on words spoken or heard during the experiment based on the following specification:

$$\begin{aligned}\text{Recalled}_{w,s,I} = \beta_0 &+ \beta_1 \textit{Total Spoken}_s + \beta_2 \textit{Total Spoken}_{s,I} + \beta_3 \textit{Total Heard}_s + \beta_4 \textit{Total Heard}_{s,I} \\ &+ \beta_5 \textit{Total Time}_s + \beta_6 \textit{Total Time}_{s,I} + \beta_7 \textit{Frequency spoken}_{w,s,O1} + \beta_8 \textit{Frequency spoken}_{w,s,I,O1} \\ &+ \beta_9 \textit{Frequency spoken}_{w,s,O2} + \beta_{10} \textit{Frequency spoken}_{w,s,I,O2} + \beta_{11} \textit{Frequency heard}_{w,s,O1} \\ &+ \beta_{12} \textit{Frequency heard}_{w,s,I,O1} + \beta_{13} \textit{Frequency heard}_{w,s,O2} + \beta_{14} \textit{Frequency heard}_{w,s,I,O2} + \varepsilon\end{aligned}$$

Here, s is a unique participant ID and I is an image number. For each word w that subject s encounters while image I is being discussed, $\text{recalled}_{w,s,I}$ is 1 in case s recalls w during the memory task for image I, and is otherwise 0. $\textit{Total Spoken}_s$ is the number of words that s speaks during the experiment, and $\textit{Total Spoken}_{s,I}$ is the number of words that s speaks when image I is being described. $\textit{Total Heard}_s$ and $\textit{Total Heard}_{s,I}$ are defined similarly for heard words; $\textit{Total Time}_s$ and $\textit{Total Time}_{s,I}$ refer to the time that s spends in the experiment and on image I, respectively. $\textit{Frequency spoken}_{w,s,O1}$ and $\textit{Frequency spoken}_{w,s,I,O1}$ are the number of times that s speaks w during the experiment and while image I is being described, respectively; $\textit{Frequency spoken}_{w,s,O2}$ and $\textit{Frequency spoken}_{w,s,I,O2}$ are defined analogously. $\textit{Frequency heard}_{w,s,O1}$, $\textit{Frequency heard}_{w,s,I,O1}$, $\textit{Frequency heard}_{w,s,O2}$, and $\textit{Frequency heard}_{w,s,I,O2}$ are defined similarly for heard words.

Model: Logit

Dependent Variable:	$\text{Recalled}_{w,s,I}$
	(1)
$\textit{Total Spoken}_s$	-0.000*
	(-2.229)
$\textit{Total Spoken}_{s,I}$	-0.009***
	(-12.622)
$\textit{Total Heard}_s$	-0.001***
	(-3.528)
$\textit{Total Heard}_{s,I}$	-0.011***
	(-15.941)
$\textit{Total Time}_s$	-0.000***
	(-19.831)
$\textit{Total Time}_{s,I}$	0.000***
	(4.353)
$\textit{Frequency spoken}_{w,s,O1}$	-0.002
	(-1.938)
$\textit{Frequency spoken}_{w,s,I,O1}$	0.326***
	(52.948)
$\textit{Frequency spoken}_{w,s,O2}$	-0.005***
	(-4.779)
$\textit{Frequency spoken}_{w,s,I,O2}$	0.350***
	(53.795)
$\textit{Frequency heard}_{w,s,O1}$	-0.006***
	(-8.023)
$\textit{Frequency heard}_{w,s,I,O1}$	0.105***
	(26.857)
$\textit{Frequency heard}_{w,s,O2}$	-0.001
	(-0.937)
$\textit{Frequency heard}_{w,s,I,O2}$	0.125***
	(34.750)
Fixed-effects	
Image FE	✓
Fit statistics	
Observations	121455
Pseudo R^2	0.168

Table 4.4. Regression of recalled (non)experimental words on total words spoken or heard.
This table reports OLS estimates of the relationship between the number of recalled words that were used during the experiment (Model 1) or that were not used during the experiment (Model 2) and the total number of words spoken and heard during the experiment. The model specifications are as follows:

$$\textit{Experiment recalled}_{s,I} = \beta_0 + \beta_1 \textit{Ttl spoken}_s + \beta_2 \textit{Ttl heard}_s + \beta_3 \textit{Ttl time}_s + \beta_4 \textit{Ttl spoken}_{s,I} + \beta_5 \textit{Ttl heard}_{s,I} + \varepsilon$$

$$\textit{Non-experiment recalled}_{s,I} = \beta_0 + \beta_1 \textit{Ttl spoken}_s + \beta_2 \textit{Ttl heard}_s + \beta_3 \textit{Ttl time}_s + \beta_4 \textit{Ttl spoken}_{s,I} + \beta_5 \textit{Ttl heard}_{s,I} + \varepsilon$$

s are all unique participant IDs and i are all $I \in \{1, 2, \ldots, 12\}$. $\textit{Ttl spoken}_s$is a count of all the words spoken by or to a subject. $\textit{Ttl spoken}_{s,I}$is a count of all the words spoken by a subject referent to a specific I. $\textit{Ttl heard}_s$is a count of all the words heard by a subject. $\textit{Ttl heard}_{s,I}$is a count of all the words heard by a subject referent to a specific I. $\textit{Ttl time}_s$is the total sum of time a subject spent in the experimental session.

Model: OLS

Dependent variables	$\textit{Experiment recalled}_{s,I}$	$\textit{Non-experiment recalled}_{s,I}$
	(1)	(2)
Constant	3.462***	7.011***
	(8.351)	(12.047)
$\textit{Ttl spoken}_s$	0.012***	0.014***
	(5.914)	(4.677)
$\textit{Ttl heard}_s$	0.001	-0.006*
	(0.520)	(-2.439)
$\textit{Ttl time}_s$	-0.001***	-0.000*
	(-5.095)	(-2.180)
$\textit{Ttl spoken}_{s,I}$	0.114***	-0.031*
	(12.959)	(-2.460)
$\textit{Ttl heard}_{s,I}$	0.035***	-0.018
	(5.189)	(-1.850)
Fixed-effects		
Image FE	✓	✓
Fit statistics		
Observations	2280	2280
R^2	0.305	0.083
Adjusted R^2	0.300	0.077

4.4.5. Founder imprinting

Hypothesis 4.5 (Founder effects and attentional perspective). *The language used by a founder predicts organizational members' cognition of the world after the founders leave the organization.*

Hypothesis 4.5 predicts that the language used by an organization's founders predicts organizational members' cognition of the world after the founder leave the organization and therefore has long-run consequences for its functioning. In order to reduce task complexity, founders use targeted language to focus their team's attention on particular features of an image. They first look at the images holistically and then, over time, establish a consensual vocabulary. We hypothesize that, because future participants in the organization O are not party to this linguistic convergence, they are taught a the founders' vocabulary in order that they can more efficiently identify the experimental images, I.

We compare the recall of founders and non-founders after they exit the experiment as follows:

1. For each experiment e and for every image I, we define the set of *Founder Words* to comprise every word used to describe I in round 1 of experiment e.
2. For each subject s in each experimental session e:
   (a) We define the joining round $J(s)$ to be the experimental round in which s joins the experiment, and we define the founder variable $f(s)$ to be 1 in case $J(s) = 1$ and 0 otherwise.
   (b) For each image I, we record the set $M(s,I)$ of *Memory Words* that s uses when describing I during the memory task.
   (c) We define the *Overlap Count* to be the size of $F(s,I) \cap M(s,I)$ or the number of Founder Words used to describe image I while s is participating in the experiment that s recalls during the Memory Task.
   (d) We calculate the percentage of Founder Words in Memory Words, $P(s,I) = \frac{\text{Overlap count}}{|M(s,I)|}$, or *Overlap Percentage*.
   (e) We count the numbers E and D of *Elements* and *Details*, respectively, that s recalls about I.

Table 4.5 presents data on *Overlap Count* and *Overlap Percentage* described above. Founder

Table 4.5. Overlap count and percentage by joining round

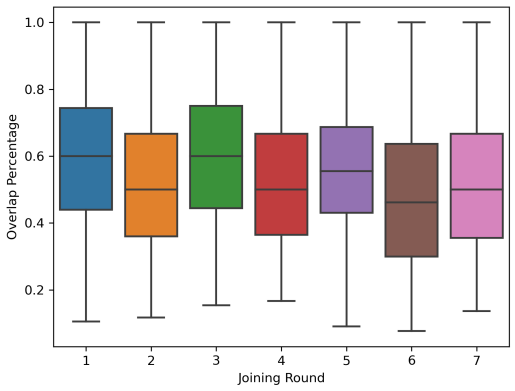
Table 4.5 presents data on the overlap between *founder words* and *memory words* by s and $I \in \{1, 2, \ldots, 12\}$. *Founder Words* is a set $F(s, I)$ of every word used by the founding team (defined as the team participating in round $R = 1$) to describe a specific image I. *Memory Words* is a set $M(s, I)$ of all the words that s uses when describing I during the memory task. The data is presented by Joining Round J in $J \in \{1, 2, \ldots, 7\}$. *Overlap Count* is calculated by taking the length of $F(s, I) \cap M(s, I)$ and averaging across all s and I for each J. *Overlap Percentage* is calculated by taking the fraction of *FounderWords* in *MemoryWords*, $P(s, I) = \frac{|F(s,I) \cap M(s,I)|}{|M(s,I)|}$ and averaging across all s and I for each J.

Joining Round	Overlap Percentage	Overlap Length
1	0.590	10.305
2	0.528	7.839
3	0.605	7.346
4	0.513	7.265
5	0.565	5.846
6	0.479	6.709
7	0.530	5.821

language is very sticky: the percentage of *Founder Words* stays stationary across all joining rounds: the average overlap percentage is 59% for subjects that join in $J = 1$ and it is 53% for subjects that join in $J = 7$. Note that the average *Overlap Count* falls from 10.305 words in $J = 1$ to 5.821 in $J = 7$; this is a mechanical result because, as the experiment progresses, the total number of words used by each organization to identify the images also falls (see Section 4.4.1). Figure 4.12 presents these results in box-and-whisker plots.

Figure 4.12. Percentage overlap by joining round

Figure 4.12 presents box-and-whisker plots on the *Percentage Overlap* between *Founder Words* and *Memory Words*, averaged across all s and $I \in \{1, 2, \ldots, 12\}$. *Founder Words* is a set $F(s, I)$ of every word used by the founding team (defined as the team participating in round $R = 1$) to describe a specific image I. *Memory Words* is a set $M(s, I)$ of all the words that s uses when describing I during the memory task. *Overlap percentage* is the fraction of *FounderWords* in *MemoryWords*, $P(s, I) = \frac{|F(s,I) \cap M(s,I)|}{|M(s,I)|}$. The data is presented by Joining Round J in $J \in \{1, 2, \ldots, 7\}$.



We know that, because organizational vocabularies converge over time, founders use more words to perform the experimental task than non-founders. We hypothesize that when a partici-

pants uses a wider vocabulary to describe an image *I*, they are likely to recall more features of the image than participants that used a more parsimonious vocabulary. Therefore, a second approach to Hypothesis 4.5 asks whether founders have better recall in the memory task than non-founders. Figure 4.13 provides evidence that they do. Every subject is asked to recall the content of the images after leaving the experiment. Founder subjects outperform non-founder subjects in identifying elements and details of the images. The difference in average times that is illustrated in Figure 4.13 has a significance level in excess of 99.9% (see Table 4.6). We hypothesise that the difference arises because founders actively engage with language creation and, as a result, are forced to pay much more attention to the overall content of the pictures; later participants can instead concentrate upon task performance and rely upon the organizational schema for language processing.

Figure 4.13. Recall by founder and non-founder members of the organization

Figure 4.13 compares recall by founders and non-founders. We counted the number of elements (e.g. tree, fence, building) and details (e.g. wispy, white, derelict) recalled by each participant. We then averaged the elements and details recalled by all founders and all non-founders. These results are displayed it box-and-whiskers plots. The top and bottom of the whiskers represent the maximum and minimum, respectively; the top of the box represents the upper quartile, the bottom of the box represents the lower quartile, and the line in the middle represents the median.

(a) Founder and non-founder recall of image elements **(b) Founder and non-founder recall of image details**

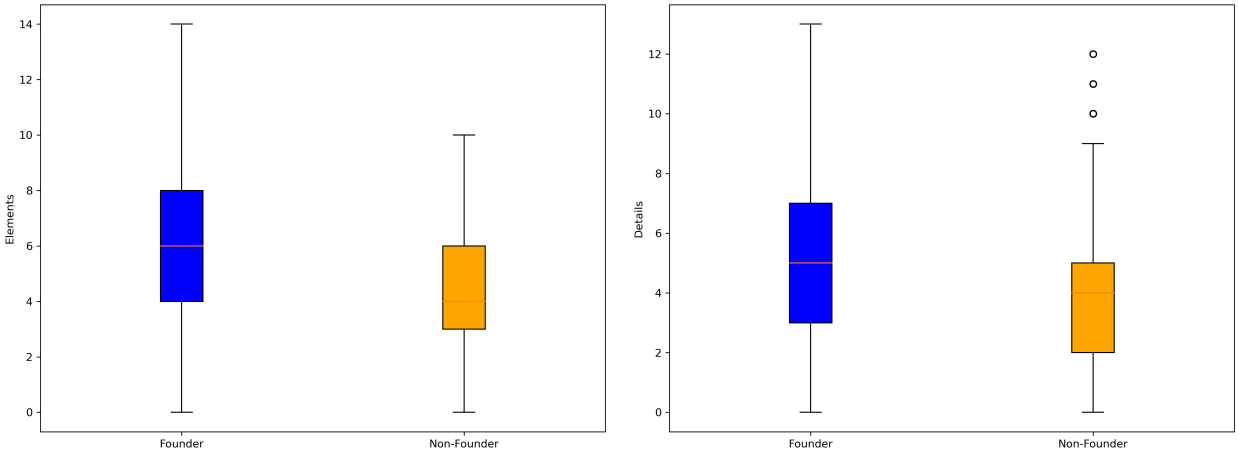


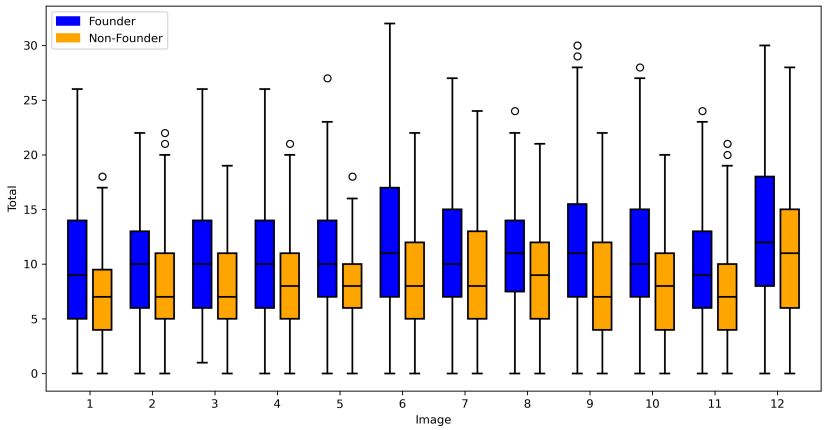
Table 4.6. Difference-in-Means of Founder and Non-Founder Recall

Table 4.6 shows the difference-in-means, t-stats, and p-values of the data displayed in Figure 9. We counted the number of elements (e.g. tree, fence, building) and details (e.g. wispy, white, derelict) recalled by each participant. We then averaged the elements and details recalled by all founder and all non-founders. "Total" adds "elements" and "details" before averaging.

	Diff-in-means
Elements	1.260***
	(10.424)
Details	1.330***
	(10.800)
Total	2.605***
	(11.123)

Figure 4.14. Recall by founder and non-founder organizational members by Image

Figure 4.14 compares recall by founders and non-founders. We counted the number of elements (e.g. tree, fence, building) and details (e.g. wispy, white, derelict) recalled by each participant. We then averaged the elements and details recalled by all founders and all non-founders. These results are displayed it box-and-whiskers plots. The top and bottom of the whiskers represent the maximum and minimum, respectively; the top of the box represents the upper quartile, the bottom of the box represents the lower quartile, and the line in the middle represents the median.



The superior founder recall illustrated by Figure 4.13 is not driven by a small number of images. Figure 4.14 shows that this result is robust across all 12 images, although performance on some images is closer than on others. Note that the variance of founder recall is higher than the variance of non-founder recall. We hypothesise that this is because non-founders rely upon organizational schema to trigger recollection and, hence, organizational schema are common to all, the quality of recall is likewise similar.

Finally, we run an OLS regression that estimates the impact of founder status, total time, total words spoken, and total words heard during the experiment on the recall of experimental words (i.e. word is in set $M(s, I)$). The results of our regression are presented in Table 4.7. The specifications

for this regression are similar to the ones presented on Table 4.4 with an added *Founder* independent variable. *Founder* is a dummy variable where 1 is assigned to founders, i.e. participants that joined in $J = 1$ of the experimental session e, and 0 is assigned to non-founders, i.e. participants that joined in $J \in \{2, 3, \ldots, 7\}$. We find *Founder* has a very significant impact on the probability of recalling a word ($p - value = 0.000$) and the magnitude of that effect is bigger than all other variables in the regression. Similarly to Table 4.4, *Ttl spoken*$_s$, *Ttl spoken*$_{s,I}$, *Ttl heard*$_{s,I}$, and *Ttl time*$_{s,I}$ are all highly significant. In short, this regression supports Hypothesis 4.5.

Table 4.7. Regression of recalled experimental words on founder-status and total words spoken or heard

This table reports OLS estimates on the impact of founder-status and total words spoken or heard during the experiment on the number of recalled words. The model specifications are as follows:

$$\begin{aligned} \textit{Experiment recalled}_{s,I} = \beta_0 + \beta_1 \text{Founder} + \beta_2 \textit{Ttl spoken}_s + \beta_3 \textit{Ttl spoken}_{s,I} + \beta_4 \textit{Ttl heard}_s + \beta_5 \textit{Ttl heard}_{s,I} \\ + \beta_6 \textit{Ttl time}_s + \beta_7 \textit{Ttl time}_{s,I} + \varepsilon \end{aligned}$$

s are all unique participant IDs and *i* are all $I \in \{1, 2, \ldots, 12\}$. *Founder* is a dummy variable where founders are tagged as 1 and non-founders tagged as 0. *Ttl spoken*$_s$is a count of all the words spoken by or to a subject. *Ttl spoken*$_{s,I}$is a count of all the words spoken by a subject referent to a specific *I*. *Ttl heard*$_s$is a count of all the words heard by a subject. *Ttl heard*$_{s,I}$is a count of all the words heard by a subject referent to a specific *I*. *Ttl time*$_s$is the total sum of time a subject spent in the experimental session and *Ttl time*$_{s,I}$is the total sum of time a subject spent exposed a specific *I*.

Model:	OLS
Dependent Variable:	*Experiment recalled*$_{s,I}$
	(1)
Founder	0.071***
	(3.840)
Ttl spoken$_s$	0.001***
	(8.290)
Ttl spoken$_{s,I}$	-0.005***
	(-8.416)
Ttl heard$_s$	-0.000
	(-1.237)
Ttl heard$_{s,I}$	-0.009***
	(-16.721)
Ttl time$_s$	-0.000***
	(-6.775)
Ttl time$_{s,I}$	0.000***
	(4.080)
Fixed-effects	
Image FE	✓
Fit statistics	
Observations	121455
Pseudo R^2	0.019

4.5. Conclusion

Language is at the heart of organizational life. It is the battleground upon which social evaluations are contested (Srivastava, Goldberg, Manian, and Potts, 2018; Ocasio, Loewenstein, and Nigam, 2015; Dunn and Jones, 2010; Mohr and Guerra-Pearson, 2010; Zhong, Loewenstein, and Murnighan, 2007; Colyvas and Powell, 2006; Kunda, 1995; Gallie, 1956); it provides the philosophical sedimentation that underpins our world views (Merleau-Ponty, 1945); and it carries the heavy metaphysical baggage that we must bear if we are to function in a world of institutions (Searle, 1995). Vocabularies therefore carry organizational histories and furnish organizational members with a cultural foundation upon which they can build meanings.

This chapter studies a particular contribution that language makes to organizations: namely, the way that vocabularies are imprinted on organizations. We present a simple model that shows how vocabularies are formed through the repeated performance of a task that requires an actor with an object to convey to others in the same organization in which category that object sits. These vocabularies are persistent and can outlast its founders. Once vocabularies start to emerge, they carry cultural understandings of meanings that serve, in turn as common ground for the future development of transactional language games. They affect how organizational members perceive environmental cues and, therefore, impact action.

It is very hard to conduct empirical work on the development and persistence of organizational vocabularies and how they impact organizational cognition. The reason is that everything that is relevant to this relationship in a real organization is endogenous. Categories are formed by, and in turn affect, vocabularies. Cognition and vocabularies are likewise bound up in one another. And the long-run evolution of vocabularies as staff turns over is very hard to trace in a complicated organization whose staff turn over in unpredictable ways and at unpredictable times.

This chapter uses an experimental set-up to address the empirical difficulties that surround studies of vocabularies and imprinting. By conducting an experiment, we are able to fix organizational processes, categories, and goals, and we can ensure that staff turnover occurs in a completely predictable way. As a result, we can investigate how processes, categories, and goals affect category formation; we can measure the evoluation of vocabularies over time; and we can measure their

long-run stability and impact on recall.

We find that organizational vocabularies converge over time and that, once they have converged, those vocabularies are fairly stable. We also find that vocabulary convergence is associated with faster task performance. However, we also show that convergence comes at a potential cost, because smaller organizational vocabularies are associated with more limited recall; that is, people in organizations with mature vocabularies attend to fewer features of their environments. In line with this finding, we demonstrate that founders, who typically have to create a working vocabulary from scratch, have broader cognitive maps of environmental cues that workers in subsequent generations. Finally, founders imprint vocabularies on organizations that significantly outlast their membership.

Our experimental approach generates clean causal statements about the impact that an organization's categories and goals have upon its vocabularies. It generates clear results about the relationship between vocabularies and imprinting, and it yields precise statements about about founder effects and the long-run impact of vocabulary choices. At the cost of some causal certainty, similar textual analysis could be used to identify corporate vocabularies in archives such as emails, board minutes, earnings call transcriptions, or press releases. The results of this type of analysis could shed light upon the ways in which attention is focussed by organizations and, hence, upon an important cultural driver of corporate-level cognition.

4.6. Implications

4.6.1. Implications for organizational theory

This chapter suggests new perspectives on important old questions in organizational theory. It sheds light on the impact of organizational vocabularies on cognition. Once formed, vocabularies have a strong effect on how members perceive external cues: they both serve as a toolkit for coordination and also inform how members cognize their environment even after they've left the organization. Our results also shed new light upon the conclusions on Weber and Camerer (2003), which concludes that post-merger organizations never reaches pre-merger performance levels. While their

paper doesn't analyse language directly, they suggest that code incompatibility could explain these results. Our results suggest the opposite: Organization 2 outperforms Organization 1 as the experiment progressed. This is because our experimental design allowed for better learning. Members entered Organization 2 as a trainee and were exposed to that organization's codes before they had to use it. By the time they had to perform the role of director, they could choose the optimal combination of Organization 1 and Organization 2 codes that would maximize task performance. This suggests that post-merger conflicts could be resolved through designing better learning processes so that members from both organizations are able to learn each other's vocabularies.

The chapter also suggests that vocabularies can be used as a microfoundation of organizational imprinting. The literature on imprinting is mostly silent on language and has instead focused on persistent meso- and macro-processes. Language offers a theoretical and empirical mechanism to better understand the dynamics of imprinting. Specifically, our experiment shows how founders can leave long-lived imprints on their organizations through vocabularies. We argue that founders create code schemas that reflect their world views and organizational members inherit those world views when they learn the languages that founders bequeath to the organization. These schemas even affect the recall of members that joined the organization after their founders had left. Members were financially incentivized to maximize their recall and, yet, their recall was greatly limited by the features highlighted by the organizational vocabulary. Finally, we shed light on the relationship between organizational and career imprints. Our experimental set-up allows us to disentangle the impact of organizational versus career imprints both in terms of organizational action (vocabulary used by participants during the experiment) and cognition (vocabulary recalled by participants after the experiment). Regarding the former, we find that organizational imprints dominate over career imprints but this difference converges over time as more participants moved from Organization 1 to Organization 2. Regarding the latter, we find that early career experiences leave a stronger imprint. Notwithstanding recency bias, participants recalled more codes from Organization 1 than Organization 2.

4.6.2. Implications for managerial practice

Our experiment suggests that cultural asynchronicity post-M&A can be reduced by instituting a more robust learning period where members of the two different organizations work together to develop a common vocabulary that uses a combination of two languages and, thus, may result in improved coordination and more efficient outcomes. Our work can also explain why organizations often fail to maximize the professional experiences of external hires. When individuals acquire experience outside an organization, this experience might not be aligned with the organization's code schema and associated categories. Therefore, individual imprints on the organization after the founding period can be slow because vocabularies are sticky.

4.A. Appendix: Tables

Table 4.A.1. Difference-In-Means of Total and Old Unigram Count

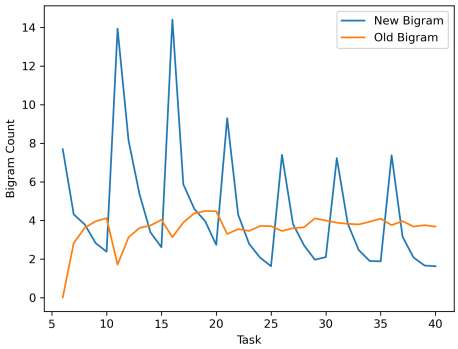
We extracted all unigrams spoken by the Director in each experiment ID, organization, task, and image combination. We then identified whether each unigram had been previously used in that organization during a previous task and calculate the number (count) of total unigrams and old unigrams. *Total unigram count* is defined as the total number of unique unigrams used per image within each task. *Old unigram count* is defined as the total number of unique unigrams used in this task that were also used in a previous task of the same organization. Subsequently, we produce an average per organization and task across all experiment ID's and image's. Given Organization 1's and Organization 2's means of size n_1 andn_2 with unknown means μ_1 and μ_2 and known standard deviations σ_1 and σ_2, we calculate the difference-in-means using a *two-sample z-statistic*. Finally, we produce a t-stat and p-value associated with each difference-in-means value. The formula used is:

$$z = \frac{(\overline{x_1} - \overline{x_2}) - (\overline{\mu_1} - \overline{\mu_2})}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

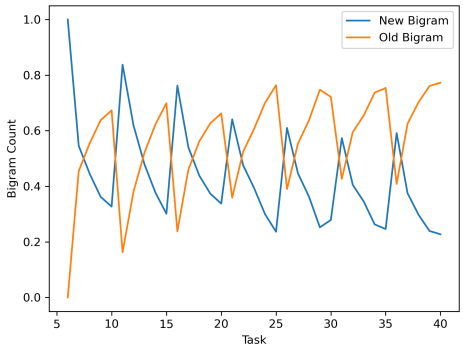
Task	Mean Org 1	Mean Org 2	Std Dev Org 1	Std Dev Org 2	t-stat	p-value
6	0.01	0.00	0.12	0.00	1.00	0.32
7	5.20	5.53	2.66	3.72	-1.10	0.27
8	6.03	6.01	3.05	3.71	0.07	0.94
9	6.12	6.48	2.97	4.40	-1.06	0.29
10	6.24	6.59	3.20	4.31	-1.02	0.31
11	4.63	4.41	2.84	2.92	0.84	0.40
12	6.55	5.60	4.71	3.18	2.68	0.01***
13	6.64	5.57	4.70	2.92	3.11	0.00***
14	6.22	5.57	4.34	3.06	1.97	0.05**
15	6.21	5.53	4.50	2.85	2.04	0.04**
16	7.57	5.20	4.67	3.05	6.76	0.00***
17	7.58	5.22	4.27	2.81	7.34	0.00***
18	7.63	5.03	4.55	2.81	7.74	0.00***
19	7.48	5.07	4.31	2.67	7.56	0.00***
20	7.03	4.79	3.66	2.82	7.64	0.00***
21	7.66	4.64	4.53	2.48	9.26	0.00***
22	6.70	5.07	3.93	2.56	5.46	0.00***
23	5.98	4.93	3.38	2.34	4.01	0.00***
24	5.88	4.89	2.92	2.28	4.13	0.00***
25	5.70	5.08	3.12	2.43	2.42	0.02**
26	7.48	4.85	4.14	2.84	8.15	0.00***
27	6.66	5.04	3.45	2.98	5.51	0.00***
28	6.21	4.94	3.08	3.14	4.44	0.00***
29	6.21	4.92	3.10	2.87	4.74	0.00***
30	6.18	4.85	3.27	2.95	4.68	0.00***
31	7.77	5.30	4.01	2.47	8.06	0.00***
32	6.85	5.67	3.59	2.47	4.12	0.00***
33	6.23	5.84	2.95	2.73	1.48	0.14
34	6.12	5.71	2.87	2.55	1.59	0.11
35	6.19	5.68	3.10	2.72	1.85	0.06*
36	7.82	4.41	4.73	2.27	9.95	0.00***
37	6.76	4.40	3.41	2.15	8.78	0.00***
38	6.01	4.33	2.99	2.31	6.60	0.00***
39	5.83	4.28	3.01	2.18	6.19	0.00***
40	5.54	3.68	2.92	1.78	8.09	0.00***

Figure 4.A.1. Bigram count and percentage by organization.
Figure 4.A.1 shows the number (count) and percentage of new and old bigrams by task. We extracted all words spoken by the Director in each experiment ID, organization, task, and image combination. We then identified whether each word was novel to an organization's vocabulary or whether it had been previously used (i.e. used in a previous task). We calculate the number of new bigrams and old bigrams and the percentage of new bigrams and old bigrams in each experiment ID, organization, task, and image combinations and then produce an average. We exclude the first 5 tasks.

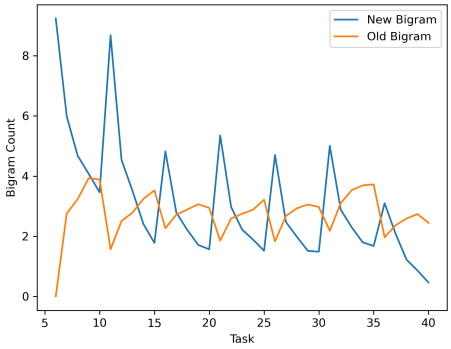
(a) Organization 1: New vs. Old Average Count



(b) Organization 1: New vs. Old Average Percentage



(c) Organization 2: New vs. Old Average Count



(d) Organization 2: New vs. Old Average Percentage

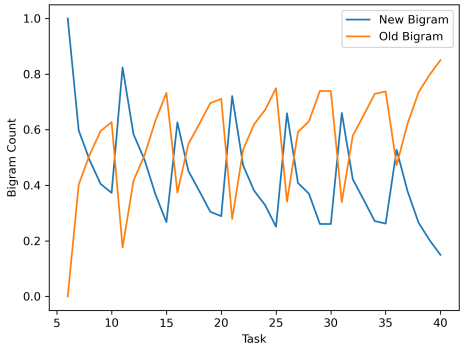
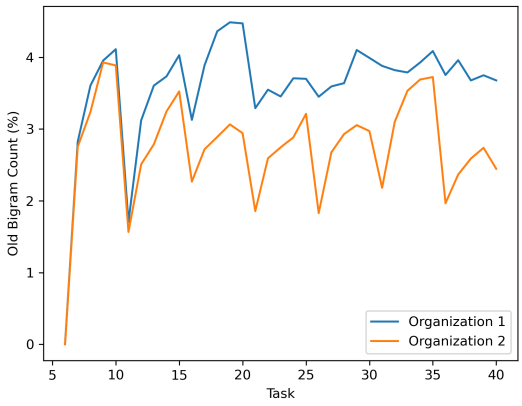


Figure 4.A.2. Old Bigram Count and Percentage.
Figure 4.A.2 shows the count and percentage of old bigrams by organization. As per Figure 1, we extracted all bigrams spoken by the Director in each experiment ID, organization, task, and image combination. We then identified whether each bigram was novel to an organization's vocabulary or whether it had been previously used (i.e. used in a previous task). We calculate the number (count) and percentage of old bigrams in each experiment ID, organization, task, and image combinations and then produce an average. We exclude the first 5 tasks.

(a) Old Bigram Count



(b) Old Bigram Percentage

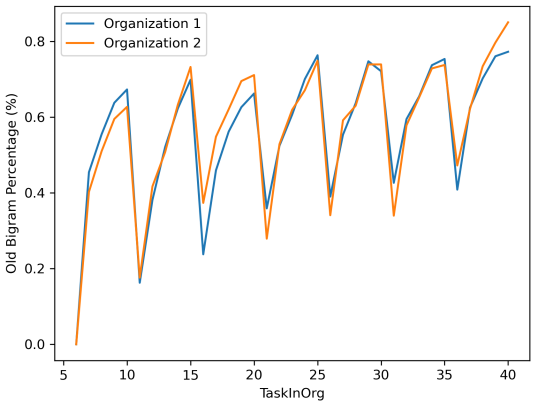
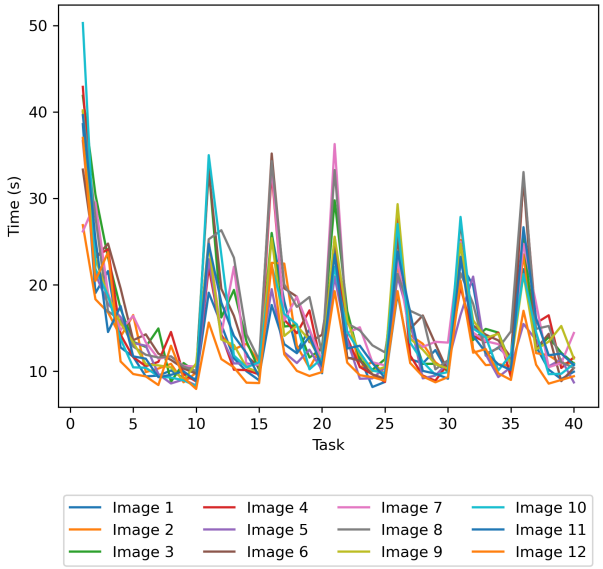


Figure 4.A.3. Time by Task per Image
Figure 4.A.3 shows the average time taken to identify each image, averaged by image. Each task is composed of 12 images. The round 1 is composed of 10 tasks and rounds 2-7 are composed of 5 tasks. At the end of each round, there is a staff turnover: the director leaves the organization, a new trainee enters the organization, and the previous trainee and associate are promoted.

(a) Time by Image in Organization 1



(b) Time by Image in Organization 2

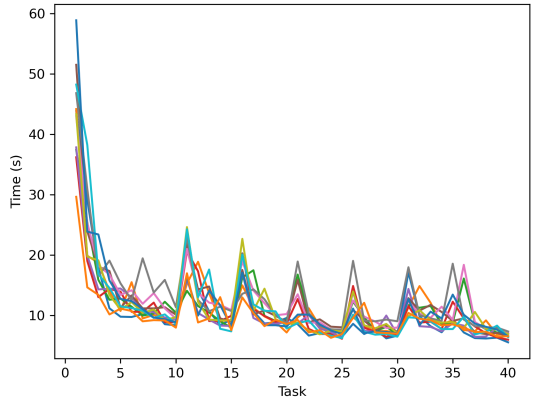


Table 4.A.2. Difference-In-Means of Average Task Time

Table 4.A.2 shows a difference-in-means test of the average task time of Organization 1 and Organization 2. During the experiment, the software automatically recorded the time taken to identify each image. Recall that each task is composed of 12 images. We add the time taken for each organization to identify the 12 images in each task and average across all experimental sessions. The Organization 2 mean is subtracted from the Organization 1 mean for each task and equivalent p-value and t-stat calculated. Given Organization 1's and Organization 2's means of size n_1 and n_2 with unknown means μ_1 and μ_2 and known standard deviations σ_1 and σ_2, we calculate the difference-in-means using a *two-sample z-statistic*. Finally, we produce a t-stat and p-value associated with each difference-in-means value. The formula used is:

$$z = \frac{(\overline{x_1} - \overline{x_2}) - (\overline{\mu_1} - \overline{\mu_2})}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

TaskInOrg	Difference-in-means	t-stat
1	-5.628*	-1.917
2	-0.569	-0.264
3	2.861**	2.247
4	0.730	0.739
5	0.526	0.680
6	-0.813	-1.227
7	-0.565	-0.611
8	-0.190	-0.284
9	-0.755	-1.194
10	0.006	0.016
11	5.771**	2.892
12	3.616**	2.378
13	2.827*	1.842
14	1.471**	2.174
15	1.179**	2.827
16	8.833***	4.517
17	4.398***	3.800
18	4.077***	4.101
19	4.009***	4.881
20	2.642***	5.746
21	13.664***	7.454
22	5.001***	6.976
23	3.764***	6.471
24	2.891***	7.561
25	2.920***	8.002
26	11.483***	8.982
27	5.467***	7.629
28	4.166***	5.574
29	2.881***	5.121
30	3.204***	8.490
31	10.547***	8.661
32	5.260***	5.574
33	3.077***	5.818
34	3.427***	5.148
35	0.603	0.618
36	14.437***	7.389
37	5.914***	8.652
38	4.707***	6.323
39	3.541***	6.418
40	4.115***	9.502

4.B. Appendix: Experiment Details

Figure 4.B.1. Participant information sheet.
This general information sheet was provided to all participants prior to their participation in the experiment. It has been reviewed and approved by both the Nuffield Centre for Experimental Social Sciences and the Saïd Business School's ethics boards.



CENTRE FOR EXPERIMENTAL SOCIAL SCIENCES

UNIVERSITY OF OXFORD

Information Sheet for LE-0066

This document is the informed consent to participate in the study "Language and Cultural Persistence" by the Centre for Experimental Social Sciences (CESS). By reading and signing this document, you declare that you fully understand the rules described below and agree to them.

Information about the experiment: The experiment in which you are participating is being conducted online using the Zoom platform and also a web browser. This experiment requires you to work in a team to perform a task. You will spend between 30 minutes and 2 hours working on the experiment. We will make an audio recording of your interactions for research purposes, but your identity will never be revealed to a third party and all research documentation will be completely anonymized, as detailed under "Confidentiality" below.

During the course of the study, Participants will be asked to rate one another's contribution to the performance of group tasks. This ranking will be used for statistical analysis and will remain anonymous. No one's ranking will be revealed to any participant in the study. At the end of the experiment, you will be asked to fill in a form that records your impressions of the experiment. Your recollections will be used for statistical analysis and they will remain confidential.

Voluntary participation and right to withdraw: Your participation in this study is entirely voluntary. This means that you agree to participate and are free to leave the laboratory at any moment without penalty. If you leave at any point, you will receive all of the payments that you have accrued to that point.

Benefits and risks: Apart from the earnings you will receive from participating, there are no other direct benefits to you from participating in this study. The results of the study may help researchers to understand how communication within organisations occurs, and how it might be more effectively organized. Participating in this study exposes you to no risks greater than those encountered in everyday life. The data collected will be anonymous.

Confidentiality: An audio recording will be made of everything that is said during the experiment for later transcription. Only a randomly assigned participant ID will identify you in the transcription. The participant ID will be assigned to the transcription and also to your questionnaire answers, but this information will be de-linked from your name and other personal information. All recordings and all other data relating to the experiment will be destroyed by the researchers after the research is published. The researchers will destroy recordings and data anyway if the research has not been published within ten years of the experiment. The Centre for Experimental Social Sciences will maintain a copy of the anonymized data, including the audio recordings, on their secure server indefinitely for the purpose of data authentication. It is possible that another participant in the experiment could record the experiment using private recording equipment at their home, but we will ask all participants refrain from this type of recording, and we will emphasise this point strongly during the experiment.

Your participation in this study will be kept as confidential as possible. The only record with your name is the administrative record that indicates that you participated and received your payment. In rare circumstances, this record may be handed over in response to a legal order. The fact that you participated in this study (but not your decisions) may also be known to people involved in the administrative process involved in the carrying out of this study. The risk of breach of confidentiality is minimal. We will ask participants not to record the experiment themselves, but we cannot check that this request is respected.

Truthfulness: During the experimental session, all information given to participants is true. This implies that we do not deceive participants. For example, if we promise you a certain payment, you will receive it and all the instructions you receive are accurate. We do not alter information that is relevant to participants. Moreover, in the case that we are responsible for an error that results in a detriment to the participants, we will inform and compensate you accordingly.

Payment: Your payment for participating in this study will be given to you via electronic funds transfer. Nobody else will know what you earned in the experiment. You will receive £3 for completing the experiment and you will be paid £8 per hour for time you spend waiting to participate in the experiment. In addition, the amount you receive for participating in the experiment will depend upon how well you perform the experimental task, as described on the instructions that you will receive during the experiment.

Contact: If you have any questions, you can contact the researchers of this study, Laura Fritsch at laura.fritsch@sbs.ox.ac.uk or Alan Morrison at alan.morrison@sbs.ox.ac.uk or CESS staff at cess@nuffield.ox.ac.uk.

Continue

Figure 4.B.2. Participant consent form.
Upon reading the participant information sheet (Figure 4.B.1 in the Appendix), all participants were given this consent form. By clicking the blue button in the bottom, they consented to participating and were moved to the detailed instructions pace (Figure 4.B.3). By closing the browser, they withdrew their consent and were exited the experiment.



Figure 4.B.3. Participant instructions.
Upon consenting to participating in the experiment (Figure 4.B.2), all participants were shown these detailed instruction sheets. By clicking continue in the second page, they were moved to the experiment's waiting room.



Instructions 1/2

You are participating in an experiment that occurs over Zoom and using a web browser. Zoom is used for communication amongst the individuals participating in today's experiment, and the web browser will be used to show pictures and to record decisions that you make during the experiment. The Zoom software used to conduct this experiment will be used to prevent all participants from switching on their webcams and sharing their screens.

Before the experiment begins, you may be asked to wait in an electronic waiting room. When you first enter the waiting room, a timer will begin that calculates the time you spend waiting to the nearest millisecond. At the end of the experiment, this number will be used to pay you in addition to the money you earn during the experiment at a rate of £8/hour. When it is your turn to participate, you will receive a notification and a link to begin the experiment.

Continue

Instructions 2/2

During the course of the experiment, you will be placed in groups of three individuals; each group in the experiment constitutes an *organization*. There are two organizations: Organization A and Organization B. One member of each organization is the *director*; another is the *associate*; and the remaining member is the *trainee*. Your role within the current organization will always appear on the top of your screen. The associate and the trainee are the organization's staff. Both members of the organization's staff can see the same screen; the director sees a different screen.

Today's experiment is comprised of a number of *Rounds*. In each round, organizations are asked to complete a number of *Tasks*. Each task is comprised of a number of iterations of a picture recognition activity. At the start of each picture recognition activity, the director's screen shows a single image while the staff's screens show thumbnails of twelve different images. Each organization's goal is to have the trainee identify the image that is shown on the director's screen as quickly as possible. To accomplish this goal, the director is asked orally to describe the image on her or his screen. The trainee uses the director's description to navigate to the correct image and then clicks on the button labelled "Accept" when it has been identified. During the game, the trainee may not speak; the associate can say anything at all to the director, and the director can respond. The trainee can use either a computer mouse or the keyboard arrow keys to navigate between the thumbnails. The trainee can press "Enter" to select a thumbnail. When a thumbnail is selected, the corresponding picture is displayed on the staff screen. When a picture is displayed, the trainee can use the left and right arrow keys to navigate to other images or can press the escape key to return to the thumbnails.

Each picture recognition activity is timed, and the goal is for the staff to identify the director's picture as quickly as possible. If the image that is on the trainee's screen matches the image on the director's screen when "Accept" is clicked, then the next picture appears on the director's screen (or the task ends). But, if the image that is on the trainee's screen does not match the image on the director's screen when "Accept" is clicked, then a message appears on their screen and a ten second penalty is applied. The director should confirm each successful match by saying 'correctly identified' whenever the picture on the screen transitions to a new image. Organization members are paid for each completed task. Each member is paid £1.00 minus the task completion time in seconds divided by 158 if the completion time is less than 158 seconds; the task payment is £0 otherwise. The task completion time is computed by the software used to run the experiment and is accurate to the nearest millisecond.

At the start of every task, every member of the organization estimates the time that the task will take and, at the end of the task, the total time elapsed is displayed on everyone's screen.

Each organization exists for one *round* of the game. The first round is comprised of 10 tasks; each subsequent round is comprised of 5 tasks.

At the end of every round:

1. Each member of the organization assigns a score to the other two members for their performance in the organization;
2. The time taken for the completed task is revealed;
3. The director of Organization B leaves the experiment;
4. Each associate is promoted to the director's position;
5. Each trainee is promoted to associate;
6. A new experimental subject enters Organization A as trainee;
7. The director of Organization A enters Organization B as trainee;
8. Every member of the organization predicts how many seconds the next task will take.

Below you can see an image representing the flow of today's experiment. Participant's in today's study begin in the waiting room (W on the left side of the image). When the experiment begins, some participant's will be randomly assignedroles in the two organizations: D for director, A for Associate and T for Trainee. Whenever an organization finishes a round, all participants shift to the right: The director of Organization 2 leaves, the associate in Organization 2 becomes the director and so on.

At the start of each round, you will be shown a screen reminding you of your role and location in the flow of the experiment. To keep the experiment moving efficiently, please refrain from communicating about the task between iterations with other participants.

Continue

Figure 4.B.4. Subject progression.
This is an illustration showing the subject's progression. The version shown here is an example of the waiting room page, which is highlighted in green. Subjects were moved to this page after reading the instructions shown in Figure 4.B.3. As subjects move through the experiment, different rectangles and circles were highlighted to indicate their progression through Organization 1 and Organization 2.

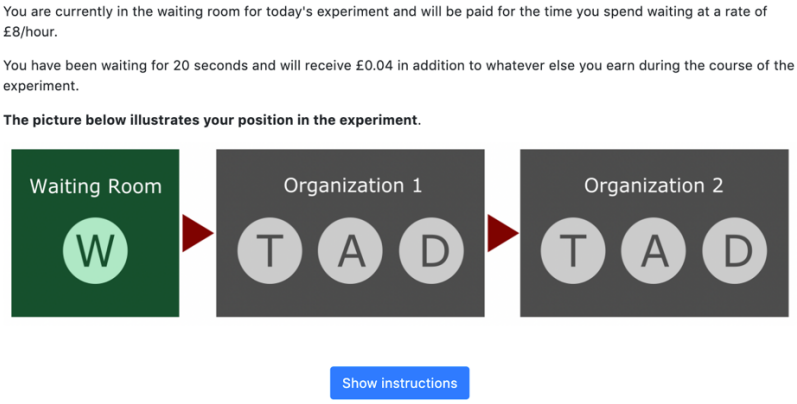


Figure 4.B.5. User interface for the director.
This is the director's user interface. For every task, they were shown each of twelve images at a random order and asked to describe the image to the Trainee and the Associate. As soon as the Trainee correctly identified the image, the image displayed in the director's user interface would change.

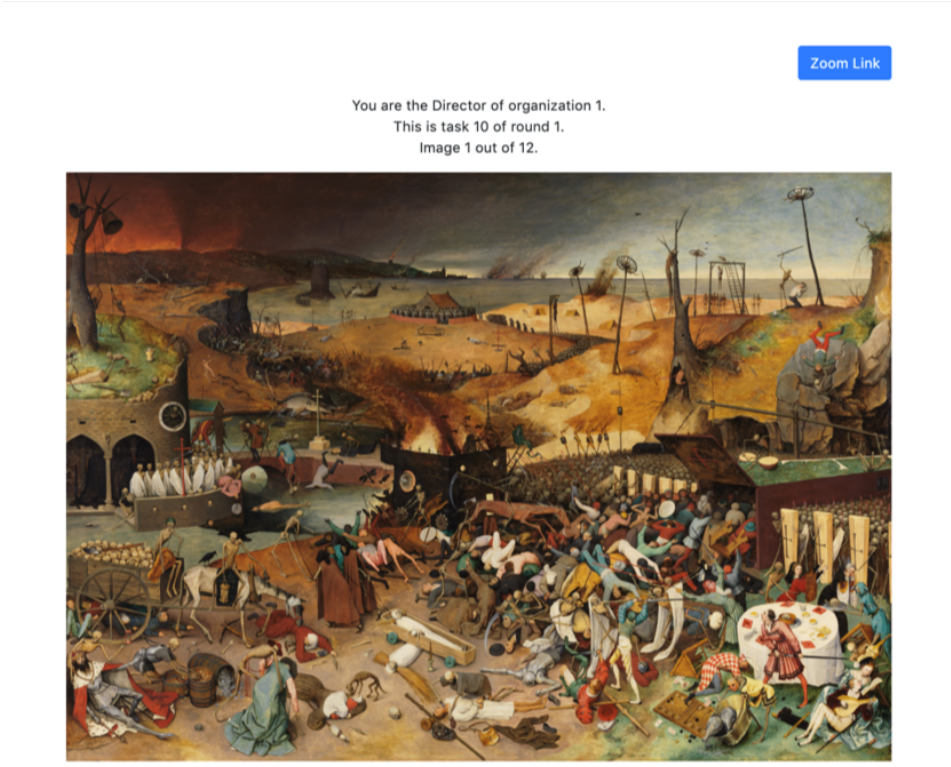


Figure 4.B.6. User interface for the trainee.
This is the director's user interface. For every task, the director would describe the image shown to them (Figure 4.B.5). The trainee and the associate had to collaborate to identify the image described as quickly as possible. They were both shown an identical grid with the twelve images. The order that the images were shown in the grid would also randomly change after every task. The trainee would choose an image by double clicking on it. If the correct image has identified, the counter on the top would change and the director would be shown a different image. If the incorrect image was identified, all three participants would receive a penalty notice (10-second deduction) and asked to try again.

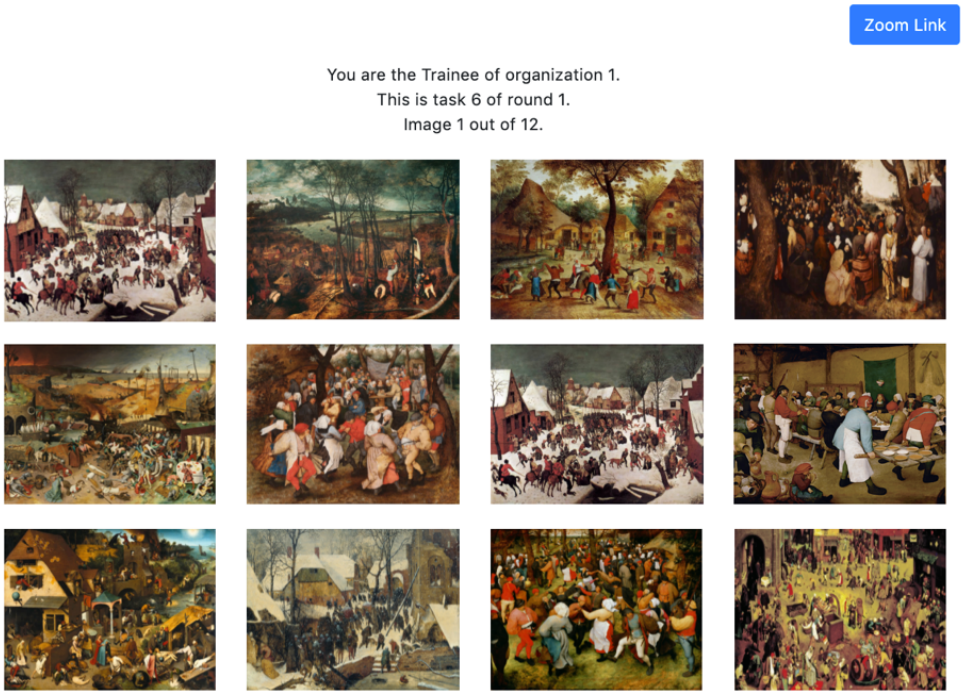


Figure 4.B.7. Sample of task pictures.
This is a random sample of the twelve pictures. Every picture was painted by the Dutch painter Pieter Bruegel the Elder. Bruegel's painting are famously complex: each of the paintings that our experimental subjects saw contained multiple identifiable features.



5

Market Reactions to Gendered Speech Patterns: Uptalk, Earnings Calls, and the #MeToo Movement

(Joint with Anant Divakaruni, Howard Jones, and Alan D. Morrison[1])

Abstract

Information is revealed by *how* people speak, as well as by what they say. Using audio recordings, we study uptalk (rising intonation) by executives in earnings calls. Unexpected uptalk by female, but not male, executives predicts lower earnings. Analysts respond to female uptalk with lower recommendations and earnings forecasts, and bid-ask spreads widen when female executives speak and use uptalk. These results are consistent with sociolinguistic studies which find that uptalk is a female-typed characteristic that signals uncertainty. #MeToo did not alter the market response to female uptalk, but it engendered a favourable response to male uptalk.

5.1. Introduction

Gender roles and expectations shape institutional arrangements. Wollstonecraft (1792) and Mill (1869) studied the effect of gender on social roles and life changes, while second- and third-wave feminist thinkers have addressed – and condemned – the gendered basis of social arrangements (Wolf, 1991; Butler, 1990; de Beauvoir, 1949; Friedan, 1963). Quantitative social scientists have

[1]The authors thank Pedro Bordalo and Mungo Wilson for their comments on an earlier draft, as well as audiences at brown-bag seminars at the University of Oxford and at the University of Bergen and reviewers at the American Economic Association Conference 2024.

identified historical-technological determinants of gendered social roles (Giuliano, 2020; Marianne, 2011). More recently, economists have started to examine the effect of gender on resource allocation (Baldiga and Coffman, 2018; Bohnet, van Geen, and Bazerman, 2016; Santacreu-Vasut, Shoham, and Gay, 2013; Alesina, Giuliano, and Nunn, 2013; Manning and Saidi, 2010; Albanesi and Olivetti, 2009; Bailey, 2006).

This paragraph falls out of the sky A growing literature examines the effects of gender on social interactions in laboratory settings (Coffman, Flikkema, and Shurchkov, 2021; Bordalo, Coffman, Gennaioli, and Shleifer, 2019; Bohnet, van Geen, and Bazerman, 2016; Reuben, Sapienza, and Zingales, 2014; Charness and Rustichini, 2011) but does not investigate these effects in real-word contexts. As well as being artificial, experiments involve one-shot engagements unable to capture long-term dynamics such as concerns about reputation.

This chapter studies the effects of gender on social interactions in which people are concerned about the reputational consequences of their behaviour. Our data are audio recordings and transcripts of quarterly earnings earnings calls for 2,993 US-listed companies between 2010 and 2019. In these calls, analysts from investors and investment banks quiz CEOs and other senior executives about their firms' current and future performance. We know the gender and job description of the participants in these calls, as well as the recommendations of stock analysts and the trading performance of the company's shares during the call.

Our analysis exploits two gendered phenomena. First, we use audio recordings of earnings calls to quantify tonal variation in speech using a variable called *uptalk*. Uptalk is a speech pattern characterized by a rising intonation in the speaker's voice at the end of a declarative utterance. [1] Uptalk has been the subject of sociolinguistic research for over 50 years (Lakoff and Lakoff, 2004) and is identified as a female-typed characteristic across different varieties of English and elsewhere. Uptalk is exhibited particularly by women (including transgender women) (Hancock, Colton, and Douglas, 2014; Hazenberg, 2013; Hazenberg, 2012) and associated with stereotypically female characteristics, both negative, e.g. uncertainty, diffidence, and lack of expertise, and positive, e.g.

[1] The term "intonation" refers to variations in a speaker's voice intensity that can convey a range of emotions such as surprise, anger, or delight. Intonation plays a critical role in communication because it not only expresses emotions but also helps convey the intended message accurately (Banziger and Scherer, 2005).

solidarity, empathy, and friendliness (Warren, 2016). In this chapter, we analyse two types of market response to uptalk in general and uptalk by women in particular: first, in analysts' recommendations and earnings forecasts following the call and, secondly, in the bid-ask spreads of the firm's share price during the earnings call itself.

The second gendered phenomenon in our dataset is the #MeToo social movement. #MeToo highlighted widespread unjust and exploitative treatment of women, especially in the workplace. The movement began at an identified time and was largely unanticipated and thus constitutes a plausibly exogenous change in social attitudes towards gender. Board-level gender diversity is a proxy for attitudes towards gender at the firm level, and an extensive literature examines its relationship with firm performance (Kim and Starks, 2016; Adams and Ferreira, 2009; Faccio, Marchica, and Mura, 2016; Adams, 2017; Adams and Funk, 2012; Matsa and Miller, 2011; Adams and Kirchmaier, 2016) but, as yet, few papers have attempted directly to relate social gender attitudes to performance. A recent exception is Lins, Roth, Servaes, and Tamayo (2019), which studies share price performance around the start of the #MeToo movement and finds that firms with proxies for a more gender-diverse culture experience significant positive abnormal returns around this date.

In the first stage of our analysis we document the distinct patterns of uptalk in our data set. Earnings calls are typically held in two consecutive phases. In the first "presentation" phase, executives make scripted remarks about the firm's recent performance and future prospects, whereas in the next "Q&A" phase they respond to analysts' questions. We find uptalk to be nearly absent in the presentation phase, but is widespread during Q&A, and significantly more so among female executives (and analysts) than in their male counterparts. We therefore focus on the Q&A part of earnings calls for the rest of our analyses. Among executives, this gendered difference shows up at the level of chief executive officer (CEO), chief financial officer (CFO), and chairperson. Female executives respond to the presence of more female executives on the call with more uptalk. By contrast, the incidence of uptalk by men decreases with the fraction of female executives present on the call. Uptalk by women increases when the firm's financial constraints are greater and decreases when analysts' recent and/or next recommendations are higher, but uptalk by male executives does not exhibit similar effects. These results are consistent with findings in the sociolinguistics literature

that uptalk is a female-typed behaviour which signals a lack of confidence.

We perform two sets of analyses to understand market responses to uptalk among executives. In the first set, we examine market reactions to uptalk over the entire sample period. In the second set, we analyse how the #MeToo movement, which began during our sample period, affected these market reactions. In both cases, we are interested in the market response to the unexpected incidence of uptalk rather than just the incidence of uptalk. We thus carry out regression analyses in two stages. In the first stage we regress the mean level of uptalk that occurs during a call against the gender composition of participating executives, as well as a range of firm and market characteristics such as firm size, recent performance, and analysts' recommendations. Since analysts must register their participation shortly before each call and signal their interest in asking questions during the call, there is quasi-exogenous variation in the number and gender composition of participating analysts which cannot be determined precisely until shortly before the call commences. Similarly, the number and gender composition of participating executives are not known to the analysts in advance, thus serving as an additional source of quasi-exogenous variation. In the second stage we test for market responses to the unexpected incidence of uptalk that is not explained by these first-stage regressions.

Across our entire sample we find that analysts make lower recommendations in response to uptalk by female executives; in particular, more unexpected female executive uptalk increases the likelihood of analyst sell recommendations and decreases the likelihood of buy recommendations. Unexpected uptalk by male executives has no such effects. The lower analyst recommendations may be a biased response to a female-typed behaviour or a rational response to an informative signal. To decide between these possibilities, we use two measures of unexpected earnings for the next quarter. The first, *Standardized Unexpected Earnings* (SUE), is a normalised measure of the extent to which the year-on-year change in earnings per share (EPS) for the firm exceeds its trend level. We find a significant negative relationship between unexpected female executive uptalk and SUE in the next quarter, and no relationship with uptalk by male executives. This suggests that analysts' lower recommendations are a rational response to unexpected female executive uptalk in that they correctly forecast the drop in earnings signalled by it. Our second measure of unexpected

earnings, *Standardized Unexpected Earnings relative to Consensus Analyst Estimates* (SUEC), is a normalized measure of the extent to which the difference between realized EPS and the consensus analyst EPS forecast differs from its trend level. We find that SUEC, which captures the accuracy of analyst EPS forecasts, is affected by the incidence of unexpected female executive uptalk in earnings calls in a way that indicates that the analysts who downgrade earnings forecasts in response to uptalk are under-reacting. This indicates that female uptalk is a signal of worse-than-expected performance in the next quarter, and that analysts respond to that signal.

We then use the width of the bid-ask spread in the firm's share price as a measure of market risk perceptions during the earnings call. We find that this spread widens by 37% on average when the responding speaker switches from a male to a female executive. Furthermore, if the corresponding difference in uptalk also increases by one standard deviation during this switch, then the spread widens by a further 4.9%: this is 13% higher than the widening that occurs when the gender of the speaker changes from a male to a female executive speaker. A possible explanation for these results is that, first, dealers make defensive quotes when a female executives starts to speak, because they anticipate that she might reveal information that lowers the firm's share price and, second, that the spreads are further widened when uptalk occurs.

While these results indicate a rational, rather than biased, market response to female uptalk, they still allow for the existence of gender bias within firms themselves. This is because an organizational procedure that we cannot capture in our dataset may assign female executives responsibility for answering questions associated with more bad news, and that this, rather than uptalk, is driving the market response. We cannot rule out this possibility entirely, but we can say that, if such bias exists, it is unlikely to affect very senior corporate executives like the CEO. When we regress bid-ask spreads on uptalk by female and male CEOs, the effect of unexpected uptalk on bid-ask spreads is qualitatively unchanged. Moreover, it seems unlikely that female CEOs impart more bad news than male CEOs. Even if this were the case, we find that uptalk is uncorrelated with the positive or negative content of the utterance, indicating that uptalk is a distinct signal from the content of the utterance itself.

The #MeToo movement, triggered by a social media post on 15 October 2017, drew worldwide

attention to gender discrimination, particularly in the workplace. The second part of our analysis acknowledges that, while market responses to gendered speech patterns may be rational, those patterns might reflect the injustice of established social structures. We investigate whether the cultural shock of #MeToo, which highlighted such injustice, affected the incidence of uptalk and the responses to it.

In the aftermath of #MeToo, uptalk by all participants in earnings calls declines significantly. There are also changes in the factors explaining uptalk. Before #MeToo, uptalk by female executives increased (and uptalk by male executives decreased) in the fraction of female executives participating in the call. After #MeToo this effect is reversed for women, suggesting that women executives become more certain in the presence of other women. The response of executive uptalk to the fraction of female analysts also changes after #MeToo. Previously, female uptalk decreased in the fraction of female analysts, while after #MeToo uptalk by female executives become indifferent to the fraction of female analysts. As for male executives, before #MeToo their uptalk level was unaffected by the fraction of female analysts, while afterwards it increased in this fraction. These results can be explained if we assume that #MeToo emboldened female analysts to ask more difficult questions, leading to more uptalk by executives of both genders.

Next we compare the market responses to uptalk before and after #MeToo. We find that #MeToo had no effect on the relationship between female executive uptalk and analyst recommendations, in that more female executive uptalk continued to lead to worse recommendations and to lower, but no less accurate, earnings forecasts. However, #MeToo brought about a change in analysts' response to male uptalk. Previously uptalk by male executives had no effect, but post-#MeToo the response by analysts to male uptalk is positive, with male executive uptalk making buy (sell) recommendations more (less) likely. This finding is consistent with the conclusion of Lins, Roth, Servaes, and Tamayo (2019) that the shares of firms that do not discriminate on grounds of gender experience positive excess returns around #MeToo. Altogether our results suggest that, following #MeToo, uptalk by female and male executives is interpreted by analysts as different signals. Female uptalk continues to be perceived as a signal of lower future short-term earnings, while male uptalk is viewed as a positive signal. In keeping with these different interpretations of female and male uptalk, while

female executive uptalk indicates lower unexpected earnings in the next quarter, male executive uptalk has no such short-term effect. Finally, we analyse the effect of #MeToo on bid-ask spreads and find that the increase in spreads in response to female uptalk continues after #MeToo and there continues to be no response to male uptalk. These results are also consistent a post-#MeToo interpretation of female executive uptalk as a marker of low confidence about the firm's immediate prospects, and of male executive uptalk as an indicator of a corporate culture in which gender discrimination is less likely.

5.2. Literature Review

5.2.1. Uptalk

In this section we summarize the literature on uptalk as a female-typed speech characteristic signalling uncertainty, on gender equality in firms, on natural language processing in finance, and on earnings conference calls as a source of data. We use Warren's definition of uptalk as "a marked rising intonation pattern found at the ends of intonation units realised on declarative utterances" (Warren, 2016, p.2), which makes declaratives sound like questions (even though not all questions have rising intonation). In particular, we identify uptalk with high tonal intensity (relative to the speaker's normal tonal intensity) at the end of the final stressed syllable of a declarative utterance. For more details, see section 5.4.

The meaning and function of uptalk have been the subject of extensive sociolinguistic research, as well as media comment, since the early 1970s (Lakoff and Lakoff, 2004). Uptalk has been interpreted under three broad headings: as a signal of a lack of confidence, as a signal of solidarity, and as a female speech characteristic.

Numerous academic works have found a strong association between uptalk and a lack of confidence (Warren, 2016; Tomlinson Jr and Tree, 2011; Barr, 2003; Spindler, 2003; Brennan and Williams, 1995; Guy and Vonwiller, 1984; Conley, O'Barr, and Lind, 1979). Many researchers include uptalk as part of a range of features typifying a powerless or low-status speech style (Loyd, Phillips, Whitson, and Thomas-Hunt, 2010; Lakoff and Lakoff, 2004; Parton, Siltanen, Hosman,

and Langenderfer, 2002; Wiley and Eskilson, 1985; Conley, O'Barr, and Lind, 1979).
Loyd, Phillips, Whitson, and Thomas-Hunt (2010) notes that people with speech characteristics signalling low status such as uptalk may not be expected to have expert knowledge, and that the expert knowledge they do possess may be 'undervalued by others' (2010: 380).

Most studies that deal with the incidence of uptalk amongst men and women agree that uptalk is particularly a feature of female speech (Warren, 2016; Daly and Warren, 2001; Henton, 1995; Henton, 1989; McConnell-Ginet, 1978). The higher incidence of uptalk in females has been observed in a wide range of English-speaking communities.[2] One of the earliest researchers to suggest that uptalk is one of a set of features that characterise women's speech was Robin Lakoff and Lakoff (2004). She argued that uptalk is a feature of 'women's language', and that 'these sorts of speech-patterns are taken to reflect something real about character and play a part in not taking a woman seriously or trusting her with any real responsibilities', since 'she can't make up her mind', and 'isn't sure of herself'. The negative associations of female uptalk have been linked to negative female stereotypes and misogyny. Levon (2016; 2020) finds that uptalk is characterised by stereotypical associations with young women and a lack of authority or credibility generally. Eckert and McConnell-Ginet (2003)[394] observe that the "construct of inarticulate female/adolescent language is popular in the media," with uptalk used as a token to justify negative views of women's speech or youth speech, and vice versa.

The association of uptalk with female speakers is supported by transgender studies. Hancock, Colton, and Douglas (2014) report that male-to-female transgender speakers who are perceived by listeners as female are more likely to use more uptalk than those who are not so perceived. These results are consistent with Hazenberg (2012; 2013), who finds that male-to-female transgender speakers embrace uptalk and that female-to-male transgender speakers tend to avoid it.

Some studies have noted an increasing incidence of uptalk among men. Lowry (2011) echoes comments by (Eckert and McConnell-Ginet, 2003). that uptalk is heard much more often than when it was commented on by Lakoff in the earlier 1970s, and that its semantic and pragmatic connotations may have changed, reflecting a change in women's position in society. A consequence

[2]For evidence on the incidence of uptalk in the United States, see Barry (2008), Ritchart and Arvaniti (2014), and Clopper and Smiljanic (2011).

of this may be more ready adoption of uptalk by male as well as female speakers. Loviglio (2008), in an analysis of gender differences in public broadcasting in the United States, likewise comments on an increase in male speakers' use of uptalk.

In this chapter, we focus on two stylised facts about uptalk which emerge from the sociolinguistics literature: first, that it is a signal of a lack of confidence in the speaker and, secondly, that it is a female-typed characteristic. The first of these stylised facts is important because uptalk as a signal of a lack of confidence in an executive may reveal something about the future performance of the firm. The second is important because we are interested in whether the market responds in different ways to the use of uptalk by female and male executives.

5.2.2. Gender and the Firm

Our analysis also bears on gender equality within firms. We discuss three relevant sub-categories of this literature, namely: (i) board diversity, (2) CEOs, and (3) Top Management Teams (TMTs). For a more comprehensive review, see Mah, Kolev, McNamara, Pan, and Devers (2023).

Board diversity

Recent studies have reported mixed results regarding the impact of gender equality at the board level. Some scholars have found that increased female board-level representation is associated with increased firm value (Carter, Simkins, and Simpson, 2003; Kim and Starks, 2016). Kim and Starks (2016) posit that this is because female directors provide complementary skills to those otherwise available to the board. Subsequent research by Carter, D'Souza, Simkins, and Simpson (2010) found positive relationship between gender diversity of boards and firm performance. Board diversity has also been found to positively impact mission fulfilment but negatively impact donations (Siciliano, 1996).

Not all research has found that greater gender equality leads to improved firm performance. Ahern and Sosyura (2015) study the effects of a Norwegian law mandating a board quota of 40% women and report a fall in share prices on the announcement of the law and a deterioration in op-

erating performance following its implementation. However, they attribute this to the appointment of younger and less experienced board members rather than to the appointment of women per se.

CEOs

Research on firms led by female executives versus male executives has likewise yielded mixed results (Mah, Kolev, McNamara, Pan, and Devers, 2023). We start by considering studies on market-based performance measures. Investors are found to react negatively to the appointment of female CEOs (Jeong and Harrison, 2017; Lee and James, 2007) and this effect is pronounced if the hire is external (Lee and James, 2007). (Dixon-Fowler, Ellstrand, and Johnson, 2013) found a negative spillover effect: other female-led organizations suffer negative market reactions following the appointment of a female CEO.

Despite initial negative reactions, long-term market performance measures (e.g. Tobin's Q and total shareholder returns) are more positive but mixed (Mah, Kolev, McNamara, Pan, and Devers, 2023). Some studies indicate outperformance by female-led firms, particularly with greater CEO discretion and in non-family-owned firms (Jeong and Harrison, 2017; Nekhili, Chakroun, and Chtioui, 2018). Conversely, a meta-analysis by Wang, Holmes Jr, Devine, and Bishoff (2018) that includes both short- and long-term measures of firm performance suggests a negative market performance effect for female-led firms.

Research on the effect of CEO gender on market performance following major organizational events has also had mixed results. Female-led firms may face negative evaluations during IPOs (Bigelow, Lundmark, McLean Parks, and Wuebker, 2014) but positive reactions during acquisitions and shareholder activist campaigns (Huang and Kisgen, 2013; Zhu, Zhu, and Ding, 2020; Francis, Hasan, Shen, and Wu, 2021). Stakeholder responses to female leadership seem situation-dependent, with less favorable reactions in positive, growth-oriented scenarios and more favorable responses in perceived risky or threatening situations, supporting the TCTF argument (Bigelow, Lundmark, McLean Parks, and Wuebker, 2014; Huang and Kisgen, 2013; Zhu, Zhu, and Ding, 2020; Francis, Hasan, Shen, and Wu, 2021).

Top Management Teams

Research on the impact of female representation in top management teams (TMTs) has also yielded mixed findings. Some studies suggest a positive link between higher female TMT representation and superior long-term market performance, especially in complex and innovative environments (Dezso and Ross, 2012; Francoeur, Labelle, and Sinclair-Desgagné, 2008; Jeong and Harrison, 2017). However, in the case of high gender pay disparity, the effect is negative (Yanadori, Kulik, and Gould, 2021). A meta-analysis finds no difference in performance between firms led by female and male CEOs, with higher performance in settings with high gender egalitarianism (Hoobler, Masterson, Nkomo, and Michel, 2018). However, research on leader gender's influence on other performance measures indicates that female CEOs receive a more favorable reaction during crises, and female CEOs of small banks show lower likelihoods of failure during crises (Palvia, Vähämaa, and Vähämaa, 2015; Cowen and Montgomery, 2020). Adhikari, Agrawal, and Malm (2019) also find that firms in which women executives' representation and compensation are greater face fewer operational lawsuits.

Of relevance to our empirical setting, Lins, Roth, Servaes, and Tamayo (2020) study the effect of the #MeToo movement on the share price of firms, using the gender mix of TMTs are a mediator. They find that firms with more women among their highest paid executives experienced significant excess returns and increased institutional ownership around the beginning of #MeToo; the effect was most pronounced in industries with few women executives or based in states with a greater gender pay gap. Similarly, Brown, Hu, Shohfi, Francis, Zhang, and Xin (2023) report that greater female participation by executives on earnings conference calls is associated with a positive market reaction and narrower bid-ask spreads.

Regarding accounting performance, meta-analyses find no direct link between CEO gender and accounting performance (Jeong and Harrison, 2017; Wang, Holmes Jr, Devine, and Bishoff, 2018). Firms led by female CEOs demonstrate higher accounting performance post-activism campaigns (Francis, Hasan, Shen, and Wu, 2021), while Zhang and Qu (2016) report worse accounting performance generally. Studies examining combined market and accounting performance show mixed results.

FInally, studies have shown that women in top positions can lead to a more women-friendly culture throughout the firm. Tate and Yang (2015) find that firms with more female leadership exhibit a smaller gender pay gap in their hiring, while Kunze and Miller (2017) report a similar effect for existing staff: greater female representation higher up in the organization leads to a narrower promotion gap between the genders.

5.2.3. Earnings calls

Most of the work done by economists using natural language processing (NLP) techniques uses text, rather than audio, data. That data is of four types: (i) news articles (Ke, Kelly, and Xiu, 2019; Loughran, McDonald, and Pragidis, 2019; Manela and Moreira, 2017; Ahern and Sosyura, 2015); (ii) financial disclosures (Brown, Crowley, and Elliott, 2020; Jiang, Lee, Martin, and Zhou, 2019; Joenväärä, Karppinen, Teo, and Tiu, 2019; Loughran and McDonald, 2014); (iii) Federal Open Market Committee minutes (Hansen, McMahon, and Prat, 2018; Boukus and Rosenberg, 2006; López-Moctezuma, 2016); and (iv) earnings conference calls (Gow, Larcker, and Zakolyukina, 2021; Hassan, Hollander, van Lent, and Tahoun, 2019; Chen, Nagar, and Schoenfeld, 2018; Larcker and Zakolyukina, 2012). A small number of studies have, like ours, also used audio data. Mayew and Venkatachalam (2012) integrate vocal markers of cognitive dissonance during earnings conference calls to detect financial misreporting. Gorodnichenko, Pham, and Talavera (2023) apply a deep learning model to detect emotions in press conferences following Federal Open Market Committee meetings and find that a positive vocal tone from Federal Reserve chairs leads to significant increases in share prices.

Earnings conference calls have been a popular source of data among researchers because transcripts are publicly available, longitudinal, and can be cleanly matched with market reactions. Earnings call data have been used to examine the impact of a range of textual features on share price, including deception (Larcker and Zakolyukina, 2012), non-answers (Gow, Larcker, and Zakolyukina, 2021), political risk (Hassan, Hollander, van Lent, and Tahoun, 2019), commonality of CEO first names (Even-Tov, Huang, and Trueman, 2020), and conversation tone (Chen, Nagar, and Schoenfeld, 2018). Of particular relevance to our analysis are recent papers on the impact of gen-

der in the context of earnings conference calls. To our knowledge, only one such paper has used textual analysis, that is Brown, Hu, Shohfi, Francis, Zhang, and Xin (2023), which investigates gender and earnings call participation. The authors find that female analysts participate less during calls and, conditional on participation, ask fewer follow-up questions. They also speak with a more positive tone and less uncertainty. However, the paper uses only textual transcripts, rather than audio files, and is silent on the conversational dynamics during the call; nor does it establish a causal relationship between gender and market reaction. Cook, Esplin, Glass, Judd, and Olsen (2019) use earnings call data, but not textual analysis, to study the impact of gender on the likelihood that management issues forecasts, the properties of those forecasts, and the resultant market reactions. They find that analysts and investors react more moderately to positive forecasts by female CEOs than male CEOs.

5.3. Data and Methodology

5.3.1. Sample Construction

This chapter uses transcripts and audio recordings of quarterly earnings calls of publicly listed US firms. We begin by collecting tickers for all US firms listed on three largest exchanges in the US (NYSE, NASDAQ, and NYSE American) and then, using those tickers, retrieve transcripts of the quarterly earnings calls convened by these firms from the Refinitiv StreetEvents database. This yields a starting sample of 60,038 earnings calls covering 3,465 publicly listed US firms from October 2010 to April 2019. The sample includes details of the name, role, and employer of 154,988 executives (CEOs, CFOs, chairpersons, and board members) and 345,659 analysts that participated in these calls. We next retrieve audio recordings of every call in the sample from EarningsCast. For each call, we map the textual and audio segments of each dialogue (presentation, question, and answer) spoken by the call participants (analysts and executives).

We generate data on the gender of call participants in the following way. First, we use the Python package *genderize* to generate a probabilistic assessment of each person's gender based on his or her forename(s). Second, we check the gender classification that *genderize* yields by

manually cross-checking each participant's name against publicly available information on them (LinkedIn profile, corporate webpages etc.). Out of the executives in our call sample, 15,707 are female and 139,281 are male; there are 34,287 female analysts and 311,372 male analysts.

We obtain data on analyst stock recommendations and earnings per share (EPS) forecasts for our sample firms from I/B/E/S. We collect both consensus mean recommendations as well as individual analyst recommendations. I/B/E/S reports recommendations on a discrete five-point scale ranging from 1 (Strong Buy) to 5 (Strong Sell); we follow Loh (2011) and reverse this coding scheme so that, in our data, 1 denotes a "Strong Sell," and 5 denotes a "Strong Buy." We map every consensus and individual recommendation issued within the 90 days following a firm's earnings call to that earnings call. Consensus recommendations are reported per month, and we therefore interpret the consensus recommendation to be the mean of all consensus recommendations issued on the firm's stock during the 90-day period following the call. We then create dummy variables *Buy* and *Sell* that identify the general direction of each consensus recommendation. *Buy* is equal to 1 in case the consensus recommendation is 5 ("Strong Buy") or 4 ("Buy"), and is 0 otherwise. Similarly, *Sell* is equal to 1 if the consensus recommendation is 1("Strong Sell") or 2 ("Sell") and is otherwise equal to 0.

In order to compute bid-ask spreads for a firm's stock while its earnings call is in progress, we use national best bid and offer (NBBO) quotes from Polygon. A stock's NBBO quote uses data from all trading venues where the stock is traded, and comprises the highest bid price and the lowest ask price for a stock. Since Polygon reports NBBO quotes every microsecond, our trading data contain nearly 2 billion NBBO quotes made during the earnings calls in our sample.

Finally, we gather data on financial metrics and stock characteristics for each firm in our sample: our financial metrics data come from Compustat and our stock data from the Center of Research in Security Prices (CRSP). From Compustat, we retrieve information on firm size (*Assets*), book-to-equity ratio (*Book/Equity*), EBITDA margin (*Profitability*), and Return on Assets (*ROA*). We also use Compustat to compute the financial constraints metric (*FinConstraints*) created by Kaplan and Zingales (1997) and later refined by Lamont, Polk, and Saaá-Requejo (2001)[3] We use the

[3] *FinConstraints* is estimated as $-1.000 \times [(IB + DP)/PPENTT - 1] + 0.283 \times [(AT + (PRCC \times CSHO) - CEQ - TXDB)/AT] + 3.139 \times [(DLTT + DLC)/(DLTT + DLC + SEQ)] - 39.368 \times [(DVC + DVP)/PPENTT - 1]$

Carhart four-factor model with CRSP data on daily closing and opening prices and stock returns to compute the volatility (*Volatility*) and cumulative abnormal returns on each firm's stock over the 89-day period preceding the call (*Momentum*).

5.3.2. Measuring Uptalk

We estimate uptalk for each earnings call as follows. First, we use the Python package *aeneas* to synchronise the speech fragments spoken by each participant in the call with the audio recording of that fragment. This process generates a *synchronization map* that identifies the time interval corresponding to each speech fragment.[4] Second, we use the Python package *pydub* to break the recording of each speech fragment into 15 second chunks. The Python package *parselmouth* is then used to extract the *sound intensity* for each of those chunks.[5] Sound intensity is measured in decibels (dB) relative to the auditory threshold pressure.[6] Finally, we compute uptalk for a given speech fragment as follows:

$$Uptalk = \begin{cases} \dfrac{Intensity_K}{\frac{1}{K-1}\sum_{k=1}^{K-1} Intensity_k} - 1, & \text{if } Intensity_K > \frac{1}{K-1}\sum_{k=1}^{K-1} Intensity_k \\ 0, & \text{otherwise} \end{cases}$$

where K denotes the total number of 15 second audio chunks in the speech fragment. The resultant uptalk measure is continuous and has a lower bound of 0.

5.3.3. Descriptive Statistics

Detailed descriptive statistics for he 1,607,428 questions and answers in our earnings call sample appear in the Appendix (tables 5.A.1a and 5.A.1a).

$-1.315 \times [CHE/ATT - 1]$, where the terms in parantheses refer to variables names in Compustat.

[4]The Appendix Table 5.A.2 presents a synchronization map for the General Motors Q1'2018 earnings call.

[5]*parselmouth* is a software wrapper for the Praat program.

[6]Praat measures the intensity of sound in air as $10 \log_{10}\left(\frac{1}{TP_0^2}\int x^2(t)dt\right)$, where $x(t)$ is the estimated sound pressure in Pascal (Pa), T is sound duration, and $P_0 = 2 \times 10^{-5} Pa$ is auditory threshold pressure.

5.4. Uptalk

This section presents statistics on the incidence of uptalk in earnings calls and, in particular, about the relationship between speaker gender, role, and uptalk. In line with the linguistics literature, we find that speeches by women contain higher levels of uptalk than those by men. Moreover, uptalk is role-dependent. For example, the gap between the level of female and male uptalk is significantly bigger for CFOs than it is for other executive roles.

Table 5.1. Uptalk by participant type and call features

This table presents OLS estimates of the relationship between uptalk levels and speech fragment characteristics in quarterly earnings calls based on the following specification:

$$Uptalk_{cpi} = \beta_0 + \beta_1 Female_{cp} + \beta_2 Executive_{cp} + \beta_3 Q\&A_{ci} + \beta_4 Y_{ci} + \lambda_c + \varepsilon_{cpi}.$$

Here, c and p are indices for the call and for each individual in the call, and i is a count of speech fragments starting from the beginning of the call. *Female* is a dummy that is equal to 1 if the call participant is female. *Executive* is a dummy that is equal to 1 if the call participant is the firm's executive, and to 0 if the participant is an analyst. *Q&A* is a dummy that is equal to 1 if the speech fragment is part of a question or answer, and to 0 if the speech fragment occurs during the presentation part of the call. The Y_{ci} are call-specific controls, and λ_c denotes call fixed effects. Standard errors are clustered by firm, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Model:	(1)	(2)	(3)	(4)
Female	0.056***	0.064***	0.064***	0.060***
	(0.015)	(0.015)	(0.015)	(0.014)
Q&A		0.084***	0.083***	0.054***
		(0.005)	(0.006)	(0.009)
Executive			-0.003	-0.005
			(0.009)	(0.009)
Call progress (%)				0.065***
				(0.013)
Call FE	✓	✓	✓	✓
Observations	2,071,052	2,071,052	2,071,052	2,039,125
Adjusted R^2	0.255	0.256	0.256	0.255

Table 5.1 reports regressions of uptalk at the call-speech level on gender, portion of call (presentation or Q&A), participant type, and call progress. In line with the linguistics literature, Table 5.1 shows that female participants exhibit significantly higher levels of uptalk than their male counterparts. There is almost no uptalk during presentations (see Figure 5.A.1b in the Appendix). Models (2)–(3) in Table 5.1 show a significantly higher level of uptalk in the Q&A part of the call; note from Figure 5.A.1b and models (2) and (3) of Table 5.1 that, while uptalk incidence varies by gender, it does not vary more according to whether a participant is an analyst or executive.

Table 5.2 shows how average levels of uptalk during the Q&A part of earnings calls vary among executives by role and gender. Female CEOs and CFOs use significantly more uptalk than their male counterparts. There is no significant gender-based difference in uptalk levels for chairpersons and board members. Note that board members in general use more uptalk in their speech than non-board members. These findings are illustrated in Figure 5.A.2 in the Appendix.

Table 5.2. Uptalk by gender and role.
This table presents OLS estimates of the relationship between uptalk levels in quarterly earnings calls by executives and executive role and gender based on the following specification:

$$Uptalk_{cpi} = \beta_0 + \beta_1 Female_{cp} + \beta_2 Executive_{cp} + \beta_3 Q\&A_{ci} + \beta_4 X_{cp} + \beta_5 Female_{cp} \times X_{cp} + \beta_6 Y_{ci} + \lambda_c + \varepsilon_{cpi}.$$

Here, c and p are indices for the call and for each individual in the call, and i is a count of speech fragments starting from the beginning of the call. *Female* is a dummy that is equal to 1 if the call participant is female. *Executive* is a dummy that is equal to 1 if the call participant is the firm's executive, and to 0 if the participant is an analyst. $Q\&A$ is a dummy that is equal to 1 if the speech fragment is part of a question or answer, and to 0 if the speech fragment occurs during the presentation part of the call. The X_{cp} and Y_{ci} are participant and call-specific controls, respectively, and λ_c denotes call fixed effects. Standard errors are clustered by firm, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Model:	(1)	(2)	(3)	(4)
Female	0.066***	0.067***	0.031*	0.033*
	(0.018)	(0.017)	(0.018)	(0.017)
Q&A	0.048***	0.048***	0.048***	0.048***
	(0.008)	(0.008)	(0.008)	(0.008)
Executive	0.337***	0.340***	0.340***	0.342***
	(0.037)	(0.038)	(0.031)	(0.032)
Call progress (%)	0.065***	0.064***	0.063***	0.063***
	(0.014)	(0.014)	(0.014)	(0.014)
CEO	0.014	0.006	0.011	0.003
	(0.012)	(0.016)	(0.013)	(0.017)
CFO	0.049***	0.051***	0.038**	0.040***
	(0.015)	(0.014)	(0.016)	(0.015)
Chairperson		0.016		0.015
		(0.018)		(0.018)
Board member		0.173		0.171
		(0.153)		(0.157)
Female × CEO			0.003	0.001
			(0.044)	(0.050)
Female × CFO			0.095**	0.092**
			(0.043)	(0.042)
Female × Chairperson				0.013
				(0.087)
Female × Board member				0.005
				(0.204)
Call FE	✓	✓	✓	✓
Observations	1,607,428	1,607,428	1,607,428	1,607,428
Adjusted R^2	0.288	0.288	0.288	0.288

In our analysis below, we will be concerned with the information that market participants derive

from executive speeches and, in order to establish the expected properties of those speeches, we regress the mean level of uptalk among participants across an entire call on the characteristics of participants and the firm conducting the call. Table 5.3 reports our results. Model (1) examines the effect of these characteristics upon the mean level of uptalk by all executives in the call. With the exception of the firm's stock price volatility in the 90 days preceding the call, none of the variables shows a significant association with mean uptalk among executives. Models (2) and (3) examine the effect of call and firm characteristics on mean uptalk levels among female executives participating in the call, and Models (4) and (5) perform the same exercise for male executives.

The table shows that the mean level of uptalk that female executives exhibit during a call is increasing in the proportion of executives participating in the call that are female, and is decreasing in the proportion of analysts that are female. Uptalk patterns for male executives is similarly increasing in the proportion of executives participating in the call that are male, and is decreasing in the proportion of analysts that are male. If uptalk is evidence of discomfort or anxiety, then this means that executives are less comfortable when their peer group has the same gender as they do. As far as we are aware, this effect has not been documented in the existing literature on uptalk; the effect is very pronounced for female executives, and is weaker for male executives.

Our interpretation of several results in Table 5.3 relies upon the generally accepted result that uptalk is a signal of a lack of confidence and is exhibited particularly by women rather than men. We combine this fact with the observation that poor corporate performance is likely a driver of lack of confidence amongst executives. That means that recent analyst recommendations positively associated with levels of executive confidence; it follows that they should also be negatively correlated with uptalk levels among female executives, but not male executives. In line with this reasoning, Model (2) demonstrates that female executive responses have less uptalk in the wake of a positive recent recommendation, and Model (4) identifies no relationship between male executive uptalk and recent analyst recommendations.

We use the same line of reasoning to explain the relationship identified in Table 5.3 between executive uptalk levels and the next analyst recommendation in Models (3) and (5). If executives know that the company's prospects are poor, then this is likely a source of a lack of confidence

Table 5.3. Factors that influence uptalk among executives in a call

This table presents OLS estimates of the relationship between the average incidence of uptalk among executives participating in quarterly earnings calls and call- and firm-level features based on the following specification:

$$Mean\ Uptalk_{cftg} = \beta_0 + \beta_1 FemaleExecs_{cft} + \beta_2 FemaleAnalysts_{cft} + \beta_3 X_{ft} + \lambda_f + \theta_t + \varepsilon_{cft}.$$

Here, c and f are indices for the call and the firm, t is the year-quarter of the call, and g is a participant grouping (e.g., all executives, female executives only, male executives only). $FemaleExecs_{cft}$ and $FemaleAnalysts_{cft}$ are the respective percentages of female executives and analysts participating in the call. X_{ft} is a vector of financial and stock characteristics and consensus analyst recommendations for firm f, and λ_f and θ_t are firm and time fixed effects, respectively. Standard errors are clustered by firm, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variables:	Mean Uptalk (Executives)	Mean Uptalk (Female Executives)		Mean Uptalk (Male Executives)	
Model:	(1)	(2)	(3)	(4)	(5)
Female Executives(%)	0.0205	0.6712***	0.6626***	-0.1533***	-0.1511***
	(0.0421)	(0.0674)	(0.0667)	(0.0230)	(0.0228)
Female Analysts(%)	0.0026	-0.2058***	-0.1996***	0.0955***	0.0941***
	(0.0263)	(0.0459)	(0.0449)	(0.0177)	(0.0175)
Log(Assets)	0.0143	-0.0095	-0.0099	-0.0095**	-0.0112**
	(0.0113)	(0.0094)	(0.0092)	(0.0045)	(0.0046)
Book/Equity	0.0054	0.0080	0.0069	-0.0016	-0.0021
	(0.0134)	(0.0086)	(0.0081)	(0.0072)	(0.0070)
Profitability	0.1988	-0.0319	-0.0133	-0.1377	-0.1696
	(0.2148)	(0.1125)	(0.1091)	(0.1490)	(0.1588)
ROA	-0.1081	-0.0357	-0.0595	0.1487	0.1264
	(0.1489)	(0.0658)	(0.0575)	(0.1024)	(0.1031)
FinConstraints	0.0001	0.0004**	0.0004**	0.0000	0.0000
	(0.0005)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Momentum(-90,-2)	0.0056	0.0009	0.0003	0.0032	0.0024
	(0.0139)	(0.0066)	(0.0066)	(0.0083)	(0.0080)
Volatility(-90,-2)	1.206**	0.0904	0.0504	0.0894	-0.0339
	(0.4959)	(0.2115)	(0.2113)	(0.3129)	(0.3206)
EPS	-0.0017	0.0001	0.0000	-0.0019	-0.0015
	(0.0029)	(0.0023)	(0.0022)	(0.0019)	(0.0018)
ΔEPS_{t-4}^{t}	0.0033	0.0017	0.0014	0.0022	0.0015
	(0.0023)	(0.0011)	(0.0011)	(0.0020)	(0.0018)
SUE		-0.0015		-0.0025	
		(0.0019)		(0.0018)	
Recent Rec (mean)		-0.0193***		0.0067	
		(0.0061)		(0.0057)	
SUE_{t+1}			-0.0021		0.0004
			(0.0014)		(0.0018)
Next Rec (mean)			-0.0144***		0.0036
			(0.0050)		(0.0048)
Firm FE	✓	✓	✓	✓	✓
Year-Q FE	✓	✓	✓	✓	✓
Observations	52,168	45,938	46,366	45,938	46,366
Adjusted R^2	0.0472	0.0807	0.0829	0.1059	0.1095

for them. The poor prospects will reduce the quality of the next analyst recommendation, and the lack of confidence will manifest in uptalk by female executives, but not male executives. In

line with this argument, Model (3) reports a significant negative association between the level of the next analyst recommendation and female executive uptalk, while Model (5) identifies no such relationship for male executives.

We have also examined the relationship between call and firm characteristics and mean uptalk among male and female analysts. We find no significant relationship between mean uptalk levels by male or female analysts and either call composition or the quality of analyst recommendations. These findings are consistent with the hypothesis that female analysts do not experience the same lack of confidence as female executives in the firm that is conducting the earnings call. These results are reported in Table 5.A.3 of the Internet appendix.

5.5. Market Reactions

This section examines the way that markets respond to the gender of executive speakers in the Q&A part of earnings calls, and to the presence of uptalk in their speech. In line with our discussion in the previous section of Table 5.3, we are concerned in this section with the information inadvertently revealed by executive speech patterns during earnings calls. If speech patterns are gendered, then information revelation must also be gendered. In order to investigate these possibilities, we first study the relationship between executive speech during an earnings call and the analyst recommendations that emerge after the call; we then investigate the way bid-ask spreads vary in real time as executive speech patterns change during an earnings call.

5.5.1. Analyst Recommendations

We start by examining the possibility that the incidence of uptalk among executives affect analyst recommendations. As reported in Table 5.3, an executive's level of uptalk depends in a predictable way upon their gender, the gender composition of participants in the earnings call, and characteristics of the executive's firm. In order to establish a causal relationship between call-level uptalk and analyst recommendations, we study the way that random variation in the level of analyst uptalk in a given call affects the recommendations that are made in the wake of the call. Our analysis proceeds

in two stages.

In the first stage, we determine how much uptalk analysts expect to encounter when a given call starts. That depends upon the group g (e.g., all executives, female executives only, male executives only) that exhibits the uptalk, upon the gender split of the executives and analysts on the call, and upon financial characteristics of the stock. Our first-stage regression is therefore as follows:

$$\textit{Predicted(Mean Uptalk)}_{cftg} = \beta_0 + \beta_1 \textit{FemaleExecs}_{cft} + \beta_2 \textit{FemaleAnalysts}_{cft} + \beta_3 X_{ft} + \lambda_f + \theta_t + \varepsilon_{cft}, \tag{5.1}$$

where c and f are indices for the call and the firm, t is the year-quarter of the call, $\textit{FemaleExecs}_{cft}$ and $\textit{FemaleAnalysts}_{cft}$ are the respective percentages of female executives and analysts participating in the call, X_{ft} is a vector of financial and stock characteristics and consensus analyst recommendations for firm f, and λ_f and θ_t are firm and time fixed effects, respectively. The results of this regression are reported in Models (2) and (4) of Table 5.3.

Once we have established the predicted level of uptalk *Predicted(Mean Uptalk)* for a given earnings call, we can compute the unanticipated uptalk $\Delta(\textit{Mean Uptalk})_{cftg}$ that occurs during the call to be the difference between the realized average level of uptalk $\textit{Mean Uptalk}_{cftg}$ and the predicted level $\textit{Predicted(Mean Uptalk)}_{cftg}$:

$$\Delta(\textit{Mean Uptalk})_{cftg} = \textit{Mean Uptalk}_{cftg} - \textit{Predicted(Mean Uptalk)}_{cftg}. \tag{5.2}$$

$\Delta(\textit{Mean Uptalk})_{cftg}$ represents random call-level variation in the level of uptalk by group g. Our second-stage regression is of consensus analyst recommendations as reported by I/B/E/S on $\Delta(\textit{Mean Uptalk})$ and a variety of control variables:

$$\textit{Outcome}_{cft} = \beta_0 + \beta_1 \Delta(\textit{Mean Uptalk})_{cftg} + \beta_2 X_{ft} + \lambda_f + \theta_t + \varepsilon_{cft},$$

where c and f are indices for the call and the firm, $\textit{Outcome}_{cft}$ is a dummy variable that captures the consensus analyst recommendation as reported by I/B/E/S, X_{ft} is a vector of financial and stock

characteristics and consensus analyst recommendations for firm f, and λ_f and θ_t are firm and time fixed effects, respectively.

The results of our second-stage regression are presented in Table 5.4. Because the regression studies the effect of unexpected variation in uptalk upon recommendations, the Table reports the results of a natural experiment and we therefore interpret the relationship that it identifies as causal. The main result in Table 5.4 is that unexpected female executive uptalk gives rise to weaker analyst recommendations. Models (1) and (2) show that unexpected uptalk by female executives results in fewer buy recommendations (I/B/E/S code 4 or 5), and Models (3) and (4) show that it results in more sell recommendations (I/B/E/S code 1 or 2).

There are two possible explanations for the effects identified in Table 5.4. First, it is possible that unexpected female executive uptalk results in weaker analyst recommendations because analysts are biased against women executives, and that, as a result of their bias, analysts respond to unexpectedly high levels of female executive uptalk with negative recommendations. Second, it is also possible that analysts have no gender bias, and that unexpected female executive uptalk occurs when the firm's earnings are likely to be lower. In the latter case, analysts might update their recommendations in response to unexpected uptalk by female executives, or later on, when the bad news presaged by this uptalk emerges.

We address the question of bias in two stages. First, we examine the relationship between unexpected uptalk and *Standardized Unexpected Earnings* (SUE) in the two quarters following the earnings call. SUE is a normalised measurement of the surprise in the change in quarterly earnings per share (EPS) from the previous year. More precisely, to compute SUE we first compute the year-on-year change in quarterly EPS each quarter. We then work out the difference between the EPS change for the current quarter and the average EPS change over the previous eight quarters. Finally, we normalise this difference by the standard deviation of EPS change over the previous eight quarters.

If the analyst response to unexpected female executive uptalk was due to gender bias, then that uptalk would have no information content and so would be unrelated to SUE. Conversely, a significant negative association between SUE and unexpected uptalk by female executives would

Table 5.4. Uptalk and analyst recommendations

This table reports logit estimates of the relationship between unexpected uptalk among earnings call participants and subsequent consensus analyst recommendations based on the following specification:

$$Outcome_{cft} = \beta_0 + \beta_1 \Delta(Mean\ Uptalk)_{cftg} + \beta_2 X_{ft} + \lambda_f + \theta_t + \varepsilon_{cft}.$$

Here, c and f are indices for the call and the firm, t is the year-quarter of the call, and g is a participant grouping (e.g., all executives, female executives only, male executives only). $Outcome_{cft}$ is a dummy variable that captures the consensus analyst recommendation as reported by I/B/E/S. In Models (1) and (2), *Outcome* is 1 in case the consensus estimate is 4 or 5 and, hence, is classified as a buy, and is otherwise 0; in Models (3) and (4), *Outcome* is 1 in case the consensus estimate is 1 or 2 and, hence is classified as a sell, and is otherwise 0. $\Delta(Mean\ Uptalk)_{cftg}$ is the unexpected level of uptalk, as defined in Equation (5.2). X_{ft} is a vector of financial and stock characteristics and consensus analyst recommendations for firm f, and λ_f and θ_t are firm and time fixed effects, respectively. Standard errors are clustered by firm, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variables:	Buy recommendation		Sell recommendation	
Model:	(1)	(2)	(3)	(4)
Δ(*Mean Uptalk*) (Female Executives)	-0.4724***	-0.4771***	0.2772*	0.3025**
	(0.1591)	(0.1741)	(0.1434)	(0.1379)
Δ(*Mean Uptalk*) (Male Executives)	-0.2238	-0.2185	0.0622	0.0629
	(0.1561)	(0.1563)	(0.0900)	(0.0885)
Δ(*Mean Uptalk*) (Female Analysts)	0.1002	0.0768	0.1009	0.1016
	(0.1961)	(0.1948)	(0.1675)	(0.1719)
Δ(*Mean Uptalk*) (Male Analysts)	-0.2142	-0.2008	-0.0573	-0.0708
	(0.2657)	(0.2706)	(0.2086)	(0.2082)
Log(Assets)	-0.2098	-0.7808***	0.3212**	0.8640***
	(0.2252)	(0.2781)	(0.1551)	(0.1987)
Book/Equity	0.6714***	0.8923***	-0.9012***	-1.195***
	(0.2022)	(0.2208)	(0.1463)	(0.1641)
Profitability	-2.956	-2.595	1.178	0.7834
	(3.719)	(3.779)	(2.701)	(2.862)
ROA	-3.305	-2.994	9.235***	8.513***
	(3.643)	(3.56)	(2.485)	(2.599)
FinConstraints	-0.0110	-0.0145*	0.0213***	0.0216***
	(0.0073)	(0.0077)	(0.0060)	(0.0062)
SUE	-0.0734	-0.0920	0.0556	0.0917**
	(0.0571)	(0.0582)	(0.0446)	(0.0452)
Momentum(-90,-2)	0.3819	0.4044	-0.2809	-0.2768
	(0.2725)	(0.2661)	(0.1802)	(0.1752)
Volatility(-90,-2)	13.68	15.5	-14.62**	-20.28**
	(11.6)	(13.27)	(6.879)	(8.494)
EPS	-0.0338	-0.0291	-0.0726	-0.0736
	(0.0807)	(0.0810)	(0.0450)	(0.0477)
ΔEPS^{t}_{t-4}	-0.0310	-0.0316	0.0081	0.0048
	(0.0707)	(0.0732)	(0.0196)	(0.0200)
Last Recommendation (buy)	5.518***	5.565***		
	(0.1496)	(0.1566)		
Last Recommendation (sell)			5.472***	5.484***
			(0.1111)	(0.1145)
Firm FE	✓	✓	✓	✓
Year-Q FE		✓		✓
Observations	9,969	9,968	17,013	17,009
Pseudo R^2	0.6699	0.6783	0.6713	0.6779

Table 5.5. Uptalk and future unexpected earnings
This table reports OLS estimates of the relationship between unexpected uptalk among earnings call participants and future unexpected earnings based on the following specification:

$$Outcome_{ft} = \beta_0 + \beta_1 \Delta(Mean\ Uptalk)_{cftg} + \beta_2 X_{ft} + \lambda_f + \theta_t + \varepsilon_{cft}.$$

Here, c and f are indices for the call and the firm, t is the year-quarter of the call, and g is a participant grouping (e.g., all executives, female executives only, male executives only). $Outcome_{ft}$ is a measure of unexpected earnings in the year-quarter after t: in Models (1)–(3), *Outcome* is equal to Standardized Unexpected Earnings (*SUE*) and in Models (4)–(6), *Outcome* is equal to Standardized Unexpected Earnings relative to Consensus analyst estimates (*SUEC*). For firm f in year-quarter t, $SUE_{ft} = (EPS_{f,t} - EPS_{f,t-4} - \mu_{ft}^{SUE})/\sigma_{ft}^{SUE}$, where μ_{ft}^{SUE} and σ_{ft}^{SUE} are the respective mean and standard deviation of $(EPS_{f,\tau} - EPS_{f,\tau-4})$ for $\tau = t-8, t-7, \ldots, t-1$ (Bernard and Thomas, 1990); and $SUEC_{ft} = (EPS_{f,t} - (Consensus\ EPS\ estimate)_{f,t} - \mu_{ft}^{SUEC})/\sigma_{ft}^{SUEC}$, where μ_{ft}^{SUEC}, and σ_{ft}^{SUEC} are the respective mean and standard deviation of $(EPS_{f,\tau} - (Consensus\ EPS\ estimate)_{f,\tau})$ for $\tau = t-8, t-7, \ldots, t-1$. X_{ft} is a vector of financial and stock characteristics and consensus analyst recommendations for firm f, and λ_f and θ_t are firm and time fixed effects, respectively. Standard errors are clustered by firm, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variables:	SUE_{t+1}			$SUEC_{t+1}$		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Δ(*Mean Uptalk*) (Female Executives)	-0.0319**	-0.0319**	-0.0329**	-0.0063**	-0.0064**	-0.0064**
	(0.0154)	(0.0155)	(0.0147)	(0.0030)	(0.0030)	(0.0030)
Δ(*Mean Uptalk*) (Male Executives)	0.0095	0.0139	0.0089	0.0013	0.0007	0.0001
	(0.0196)	(0.0202)	(0.0188)	(0.0042)	(0.0043)	(0.0043)
Δ(*Mean Uptalk*) (Female Analysts)		0.0307	0.0111		-0.0005	-0.0005
		(0.0358)	(0.0340)		(0.0069)	(0.0070)
Δ(*Mean Uptalk*) (Male Analysts)		-0.0486	-0.0462		0.0066	0.0064
		(0.0351)	(0.0314)		(0.0072)	(0.0073)
Last Recommendation	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Year-Q FE	✓	✓	✓	✓	✓	✓
Firm/Stock controls			✓			✓
Observations	37,944	37,944	37,944	37,970	37,970	37,970
Adjusted R^2	0.045	0.045	0.122	0.026	0.026	0.028
Mean (SD) of Dependent Variable		-0.08 (1.03)			-0.02 (0.34)	

suggest that the analyst reactions reported in Table 5.4 were likely rational. In order to conduct our first test of gender bias in analyst recommendations, we therefore regress next-period SUE on unexpected uptalk. Our results are reported in Models (1), (2), and (3) of Table 5.5. The table reports a significant negative relationship between unexpected uptalk by female executives and SUE for the quarter immediately after the earnings call, and no such relationship for male analysts. We conclude that the analyst recommendations made in the wake of unexpected uptalk by female executives are likely rational; the absence of any analyst response to unexpected uptalk by male executives is likewise rational.

While the regressions of Models (1), (2), and (3) in Table 5.5 suggest that analyst recommenda-

tions move in the right *direction* after unexpected female executive uptalk, those regressions cannot establish that the *scale* of the reaction is accurate. We investigate whether the difference between realised earnings and consensus analyst earnings estimates changes in the wake of unexpected female executive uptalk. Our measure of this difference is *Standardized Unexpected Earnings Relative to Consensus Analyst Estimates* (SUEC). In order to compute SUEC, we measure the difference between the realised and consensus EPS estimate for the firm every quarter; for clarity, call this *diff*. We then subtract the average difference between realised and consensus EPS estimates for the firm over the preceding eight quarters from *diff* to obtain *net diff*. Lastly, we normalize *net diff* by the standard deviation of *diff* over the past eight quarters.

If the accuracy of consensus analyst earnings was unaffected by unexpected female executive uptalk, then there should be no relationship between SUEC and unexpected female executive uptalk. Models (4), (5), and (6) of Table 5.5 indicate that this is not the case: there is a negative and significant relationship between SUEC and $\Delta(\textit{Mean Uptalk})(\text{Female Executives})$. This means that, on average, analyst earnings estimates fail to move far enough in response to unexpected female executive uptalk. While this result suggests that the market does not respond completely rationally to the information in female executive speech patterns, the negative coefficient on $\Delta(\textit{Mean Uptalk})(\text{Female Executives})$ in Models (4), (5), and (6) is strong *prima facie* evidence against gender bias, because it suggests that earnings estimates should be revised down further than they currently are.

The results in Table 5.5 suggest a rational relationship between unexpected female executive uptalk and analyst recommendations, but they leave unanswered the question of whether analysts understand this relationship. It is possible that unexpected female executive uptalk occurs because the speaker anticipates bad news about her firm, and that analyst recommendations change only when that news is revealed. It is also possible that analysts respond directly to unexpected uptalk, and that the associated news arrives after their recommendations change. We distinguish between these possibilities in two ways.

First, we regress the average analyst recommendation as reported by I/B/E/S in each of three periods after the earnings call on unexpected uptalk by male and female executives. Our regression

results are reported in the Appendix Table 5.A.4. The negative response of analyst recommendations to unexpected female executive uptalk occurs only in the first three weeks after the call. This evidence is consistent with the hypothesis that analysts understand that this type of uptalk is an advance indicator of corporate problems, and that they update their forecasts ahead of the arrival of precise information.

Second, we perform a more granular investigation of the speed with which buy and sell recommendations emerge after unexpected uptalk. In order to do so, we regress a dummy variable that indicates whether the average I/B/E/S recommendation in week $w = 1, 2, \ldots, 6$ after an earnings call is a buy or a sell recommendation against unexpected uptalk ($\Delta(\textit{Mean Uptalk})$) by female and male executives and analysts, as well as a variety of control variables. The $\Delta(\textit{Mean Uptalk})$ coefficient results for male and female executives are illustrated in Figure 5.A.4, which appears in the Appendix.

Panels (a) and (b) of Figure 5.A.4 show coefficients for sell recommendation dummies in the weeks after the earnings calls for $\Delta(\textit{Mean Uptalk})$ by female and male executives, respectively. In line with the results of Table 5.4, only $\Delta(\textit{Mean Uptalk})$ (Female Executives) has a significant impact upon analyst recommendations; that effect is negative and occurs immediately after the earnings call. Completely analogous results are reported in Panels (c) and (d) for buy recommendations. Taken together, Panels (a) and (c) suggest that analysts respond to unexpected female executive uptalk and not to the news that it anticipates. Panels (e) and (f) present week-by-week coefficients for the difference between the median analyst recommendation per week and the consensus recommendation over the 90-day period after the call. While there is no significant difference for $\Delta(\textit{Mean Uptalk})$ (Male Executives) in any week after the earnings call in which it occurs, $\Delta(\textit{Mean Uptalk})$ (Female Executives) is associated with very negative recommendations immediately after the call. Once again, these data are consistent with the hypothesis that at least some analysts understand, and respond to, the information content of $\Delta(\textit{Mean Uptalk})$ (Female Executives).

We have now investigated the relationship between uptalk during an earnings call and the analyst recommendations and realised earnings that emerge after the call. We find no significant

relationship between either of these post-call data and unexpected male executive uptalk. In contrast, unexpected uptalk by female executives is significantly associated with lower analyst recommendations. Moreover, it appears that those recommendations are rational: realised earnings after unexpected female executive uptalk are lower than usual. We have also presented evidence that suggests that analysts understand this relationship: the significant effects that we identify emerge soon after earnings calls and so likely pre-date the release of the information that is anticipated in female executive uptalk.

One way to gain additional insight into the question of whether or not market participants understand the information content of female executive uptalk is to investigate the real-time effect of that uptalk upon market prices. The following section undertakes that investigation.

5.5.2. Risk Perceptions

This section examines the real-time effect that unexpected executive uptalk during an earnings call has upon bid-ask spreads. We know from Section 5.5.1 that unexpected uptalk by female executives is associated with lower analyst recommendations and with weaker realised earnings and, hence, that this type of uptalk is a signal of potential bad news. Market awareness of this fact could manifest in two ways. First, if a female executive starts to use uptalk, then she indicates that there is downside risk in the stock and so generates a heightened perception of risk that results in wider bid-ask spreads. Second, when a female executive starts to speak for the first time, market participants face the higher-order risk that she might start to use uptalk. If that risk is significant, then it should cause bid-ask spreads to widen as soon as a female executive starts to speak.

We use NBBO quotes from Polygon to investigate these two potential sources of risk. At every microsecond t during an earnings call, we define $\textit{Percentage Spread}_t = (\textit{Ask} - \textit{Bid})/((\textit{Ask} + \textit{Bid})/2)$ to be the bid-ask spread of the stock price quote expressed as a percentage of its midpoint; *Percentage Spread* therefore ranges between 0 and 1. We then define *Quoted Spread* for each executive response in the call to be the arithmetic average of every *Percentage Spread* calculated

during the response:

$$Quoted\ Spread(Response) = \frac{\sum_{Every\ microsecond\ t\ in\ Response} Percentage\ Spread_t}{Number\ of\ microseconds\ in\ Response}. \tag{5.3}$$

Bid-Ask Spreads and Female Executive Uptalk

We now investigate the contemporaneous effect of female executive uptalk on market risk perceptions. In order to do so, we examine the real-time changes in *Quoted Spread* that occur when female executives use uptalk. Our measure of unexpected uptalk, $\Delta(Mean\ Uptalk)$, is computed over the length of an entire call and, hence, cannot yield tick-by-tick data. We therefore use the raw measure of uptalk in our regressions.

Table 5.6 reports regressions of the *Quoted Spread* for executive responses on executive uptalk interacted with the gender of the executive speaking and other variables. In line with our hypothesis that female executive uptalk generates heightened risk perception, we find that a unit increase in female executive uptalk is associated with a three percentage point increase in *Quoted Spread*. In contrast, uptalk in male executive responses serves marginally to reduce the mean bid-ask spread. The average value of *Quoted Spread* is nine percentage points wider when female executives respond to female analysts, and is unrelated to uptalk.

On average, bid-ask spreads narrow by 69% over the course of the earnings call. We hypothesise that this effect obtains because, as information is revealed over the course of an earnings call, market perceptions of risk reduce. Moreover, longer responses by executives correspond to higher bid-ask spreads: every doubling of the response length (in minutes) corresponds to a seven percentage point increase in *Quoted Spread* from the sample average.

Our evidence that unfavourable market reactions to female executive uptalk is rational does not rule out the possibility of bias in financial markets. We have argued that the market reactions are rational, because it seems likely that female executives reveal bad news about their firms when they use uptalk. But it is possible that the assignment of executives to questions is not random: if female executives are consistently forced to talk in situations where there is non-public bad news about their firms, then, in responding to their uptalk, market participants could simply be capturing

Table 5.6. Uptalk and bid-ask spreads
This table presents GLM logit estimates of the relationship between uptalk incidence during executive responses to analyst questions and the mean bid-ask spread of the firm's stock during a given response, based on the following regression:

$$\begin{aligned} \textit{Quoted Spread}_{cftpi} = {} & \beta_0 + \beta_1 \textit{Uptalk}_{cftpi} + \beta_2 \textit{Female Executive}_{cftpi} + \beta_3 \textit{Female Analyst}_{cfti} \\ & + \beta_4 \textit{Uptalk}_{cftpi} \times \textit{Female Executive}_{cftpi} + \beta_5 \textit{Uptalk}_{cftpi} \times \textit{Female Analyst}_{cfti} \\ & + \beta_6 \textit{Female Executive}_{cftpi} \times \textit{Female Analyst}_{cfti} \\ & + \beta_7 \textit{Uptalk}_{cftpi} \times \textit{Female Executive}_{cftpi} \times \textit{Female Analyst}_{cfti} + \beta_8 X_{cfti} + \lambda_c + \theta_p + \varepsilon_{cftpi}. \end{aligned}$$

Here, c, f, and p are indices for the call, the firm, and the responding executive, respectively, and i counts question-response pairs from the start of the call. $\textit{Quoted Spread}_{cftpi}$ is the average percentage bid-ask spread for response i by executive p in call c for firm f in year-quarter t; it is calculated using the formula in Equation (5.3). $\textit{Uptalk}_{cftpi}$ is the average level of uptalk during the same response. *Female Analyst* and *Female Executive* are dummies that are equal to 1 in the respective cases where a question is posed by a female analyst and answered by a female executive, and that are otherwise equal to 0. X_{cfti} is a vector, which includes *Call Progress* and *Answer Duration*: $\textit{Call Progress}_{cfti}$ is the percentage of the earnings call that elapsed before response i in call c, expressed as a number between 0 and 100; $\textit{Answer Duration}_{cfti}$ is the duration in minutes of response i in call c. Standard errors are double-clustered by firm and date of the call, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variable:	*Quoted Spread*		
Model:	(1)	(2)	(3)
Uptalk	-0.0056*	-0.0057*	-0.0068**
	(0.0028)	(0.0028)	(0.0030)
Uptalk × *Female Executive*	0.0299**	0.0300**	0.0311**
	(0.0129)	(0.0130)	(0.0153)
Female Analyst		0.0023	0.0001
		(0.0125)	(0.0136)
Uptalk × *Female Analyst*			0.0134
			(0.0109)
Female Executive × *Female Analyst*			0.0935**
			(0.0456)
Uptalk × *Female Executive* × *Female Analyst*			-0.0169
			(0.0542)
Call Progress (%)	-0.0069***	-0.0069***	-0.0069***
	(0.0009)	(0.0009)	(0.0009)
Log (*Answer Duration*)	0.0669***	0.0667***	0.0667***
	(0.0049)	(0.0049)	(0.0049)
Call FE	✓	✓	✓
Executive FE	✓	✓	✓
Observations	387,930	387,772	387,772
Pseudo R^2	0.775	0.775	0.776

a form of within-firm bias.

In order to investigate the possibility that female executives are forced to talk in situations where bad news might be revealed, we categorize executive speech fragments according to the executive's role, and we run the regressions of Table 5.6 separately for CEOs, CFOs, Chairpersons, and Board Members. Our results are reported in the Appendix in Table 5.A.11. Female executive uptalk has

a significant effect upon *Quoted Spread* only when the executive is a CEO. Moreover, uptalk by male CEOs narrows the bid-ask spreads, although the economic magnitude of this effect is much lower.

Because CEOs are the most important executives who participate in earnings calls, we view it as implausible that women executives are deliberately being assigned the job of revealing bad news to the market. Because CEO uptalk is the only phenomenon that appears to affect market risk perceptions, we conclude that CEOs are more likely to be called upon to speak about matters that render bad news salient. That is, the results in Table 5.A.11 are consistent with the hypothesis that the responsibility for talking about potentially damaging stories is assigned by role, and not by gender.

Bid-Ask Spreads and Female Executive Participation

We now investigate the possibility that market participants have an increased perception of risk when a female executive starts to speak for the first time during the call. That this is a possibility because a female executive may use uptalk and, as noted above, female executive uptalk is viewed as a risky signal.

In order to study whether introducing a female speaker increases market risk perceptions, we measure the within-earnings-call dynamics of bid-ask spreads in the following way. First, we list the question/answer combinations in the Q&A section of the call in the order in which they occur. Second, we compute the average bid-ask spread for each executive response to a question. Third, for every consecutive pair of responses, we record the percentage change $\Delta(\textit{Quoted Spread})$ in the average spread, the change in uptalk $\Delta \textit{Uptalk}$, and the change in the executive speaker's gender $\Delta(\textit{Exec Gender})$. We then regress $\Delta(\textit{Quoted Spread})$ on dummies for the values of $\Delta(\textit{Exec Gender})$, on $\Delta \textit{Uptalk}$, and on other variables. Table 5.7 reports our results.

When a female executive succeeds a male executive during the Q&A part of the earnings call, there is a corresponding 37% increase in *Quoted Spread*. Other changes in executive gender do not have a significant effect on the bid-ask spread. These results are consistent with the hypothesis that market perceptions of risk increase when a female executive starts to speak.

Table 5.7. The effect of executive gender dynamics and uptalk on bid-ask spreads

This table reports OLS estimates of the within-earnings-call dynamics of bid-ask spreads on genders of the responding executives, occurrence of uptalk in their responses, and call progress, based on the following regression:

$$\Delta(\textit{Quoted Spread})_{cft,n} = \beta_0 + \beta_1 \Delta(\textit{Exec Gender})_{cft,n} + \beta_2 \Delta \textit{Uptalk}_{cft,n} + \beta_3 \Delta(\textit{Exec Gender})_{cft,n} \times \Delta \textit{Uptalk}_{cft,n} + \beta_4 \textit{CallProgress}_{cft,n} + \beta_5 \Delta(\textit{Exec Gender})_{cft,n} \times \textit{Call Progress}_{cft,n} + \beta_6 X_{cft,n} + \lambda_c + \varepsilon_{cft}.$$

Here, c and f are indices for the call and the firm, t is the year-quarter of the call, and n counts the transitions between executive speakers in call c for firm f in year-quarter t. As described in the text, $\Delta(\textit{Quoted Spread})_n$ is the percentage change in the average bid-ask spread for speaker change n in an earnings call, $\Delta(\textit{Exec Gender})_n$ is the corresponding change in speaker gender, $\Delta \textit{Uptalk}_n$ is the change in uptalk, and $\textit{Call Progress}_n$ is the percentage of the earnings call that has elapsed when speaker change n occurs. $X_{cft,n}$ is a vector of controls that includes $\textit{Answer Duration}_n$ measured in minutes and a $\textit{Female Analyst}_n$ dummy that is equal to 1 if the question in between the speeches at index n is posed by a female analyst and is otherwise 0. Standard errors are double-clustered by firm and date of the call, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variable:	$\Delta(\textit{Quoted Spread})$		
Model:	(1)	(2)	(3)
$I(\Delta(\textit{Exec Gender}) = \text{M to F})$	0.3702***	0.3683***	1.553***
	(0.0757)	(0.0741)	(0.3535)
$I(\Delta(\textit{Exec Gender}) = \text{F to M})$	0.0707	0.0761	0.4257*
	(0.0520)	(0.0546)	(0.2196)
$I(\Delta(\textit{Exec Gender}) = \text{F to F})$	-0.0291	-0.0311	-0.3800**
	(0.0457)	(0.0470)	(0.1538)
$\Delta \textit{Uptalk}$		0.0294***	0.0294***
		(0.0106)	(0.0106)
$I(\Delta(\textit{Exec Gender}) = \text{M to F}) \times \Delta \textit{Uptalk}$		-0.0001	-0.0097
		(0.0861)	(0.0858)
$I(\Delta(\textit{Exec Gender}) = \text{F to M}) \times \Delta \textit{Uptalk}$		-0.0267	-0.0249
		(0.0433)	(0.0429)
$I(\Delta(\textit{Exec Gender}) = \text{F to F}) \times \Delta \textit{Uptalk}$		0.0156	0.0149
		(0.0515)	(0.0515)
Call Progress (%)	-0.0038***	-0.0038***	-0.0034***
	(0.0006)	(0.0006)	(0.0006)
$I(\Delta(\textit{Exec Gender}) = \text{M to F}) \times$ *Call Progress (%)*			-0.0165***
			(0.0042)
$I(\Delta(\textit{Exec Gender}) = \text{F to M}) \times$ *Call Progress (%)*			-0.0050*
			(0.0026)
$I(\Delta(\textit{Exec Gender}) = \text{F to F}) \times$ *Call Progress (%)*			0.0047**
			(0.0020)
Log (*Answer Duration*)	0.0098	0.0091	0.0089
	(0.0157)	(0.0157)	(0.0157)
Female Analyst	-0.0472	-0.0472	-0.0477
	(0.0290)	(0.0290)	(0.0290)
Call FE	✓	✓	✓
Observations	332,835	332,835	332,835
Adjusted R^2	0.114	0.114	0.114
Mean (SD) of $\Delta(\textit{Quoted Spread})$ when $\Delta(\textit{Exec Gender})$ = M to M		0.38 (4.49)	

If the reason that female executive speakers increase risk perceptions is that uptalk by those executives may reveal bad news, then female executive participation later in the call is likely to

Table 5.8. The effect of first-time executive gender switches and uptalk on bid-ask spreads
This table reports OLS estimates of the within-earnings-call dynamics of bid-ask spreads on genders of the responding executives, occurrence of uptalk in their responses, and call progress, based on the following regression:

$$\Delta Quoted\ Spread_{cft,n} = \beta_0 + \beta_1 \Delta(Exec\ Gender)_{cft,ns} + \beta_2 \Delta Uptalk_{cft,n} + \beta_3 \Delta(Exec\ Gender)_{cft,ns} \times \Delta Uptalk_{cft,n} + \beta_4 Call\ Progress_{cft,n} + \beta_5 \Delta(Exec\ Gender)_{cft,ns} \times Call\ Progress_{cft,n} + \beta_6 X_{cft} + \lambda_c + \varepsilon_{cft,n}.$$

Here, c and f are indices for the call and the firm, t is the year-quarter of the call, and n counts the transitions between executive speakers in call c for firm f in year-quarter t. As described in the text, $\Delta(Quoted\ Spread)_n$ is the percentage change in the average bid-ask spread for speaker change n in an earnings call. For $s \in \{M\ to\ F, F\ to\ M\}$, $\Delta(Exec\ Gender)_{sn}$ is a dummy that is equal to 1 if gender change s occurs for the first time during the call at speaker change n, and is otherwise equal to 0. $\Delta Uptalk_n$ is the change in uptalk at speaker change n, and $Call\ Progress_n$ is the percentage of the earnings call that has elapsed when change n occurs. $X_{cft,n}$ is a vector of controls that includes $Answer\ Duration_n$ measured in minutes and a $Female\ Analyst_n$ dummy that is equal to 1 if the question in between the speeches at index n is posed by a female analyst and is otherwise 0. Standard errors are double-clustered by firm and date of the call, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variable:	$\Delta(Quoted\ Spread)$			
Model:	(1)	(2)	(3)	(4)
$I(\Delta(Exec\ Gender) = \text{M to F})_{\text{first}}$	0.7403***	1.693***		
	(0.1430)	(0.4892)		
$I(\Delta(Exec\ Gender) = \text{F to M})_{\text{first}}$			0.1937	0.8650**
			(0.1183)	(0.3587)
$\Delta Uptalk$		0.0259***		0.0306***
		(0.0099)		(0.0103)
Call Progress (%)	-0.0035***	-0.0033***	-0.0037***	-0.0036***
	(0.0006)	(0.0006)	(0.0006)	(0.0006)
$I(\Delta(Exec\ Gender) = \text{M to F})_{\text{first}} \times \Delta Uptalk$		0.2524		
		(0.2224)		
$I(\Delta(Exec\ Gender) = \text{M to F})_{\text{first}} \times Call\ Progress\ (\%)$		-0.0168**		
		(0.0069)		
$I(\Delta(Exec\ Gender) = \text{F to M})_{\text{first}} \times \Delta Uptalk$				-0.0865
				(0.1108)
$I(\Delta(Exec\ Gender) = \text{F to M})_{\text{first}} \times Call\ Progress\ (\%)$				-0.0108**
				(0.0049)
Log (*Answer Duration*)	0.0109	0.0101	0.0125	0.0117
	(0.0156)	(0.0156)	(0.0156)	(0.0156)
Female Analyst	-0.0460	-0.0466	-0.0466	-0.0466
	(0.0289)	(0.0290)	(0.0289)	(0.0289)
Call FE	✓	✓	✓	✓
Observations	333,990	333,990	333,990	333,990
Adjusted R^2	0.114	0.115	0.114	0.114

have a smaller effect on risk perceptions. The reason is that bad news is more likely already to have been revealed later earlier in the call, by which time several speeches by female executives might have already occurred. We present three pieces of evidence that are consistent with this claim.

First, Model (3) in Table 5.7 shows a very significant and negative interaction between

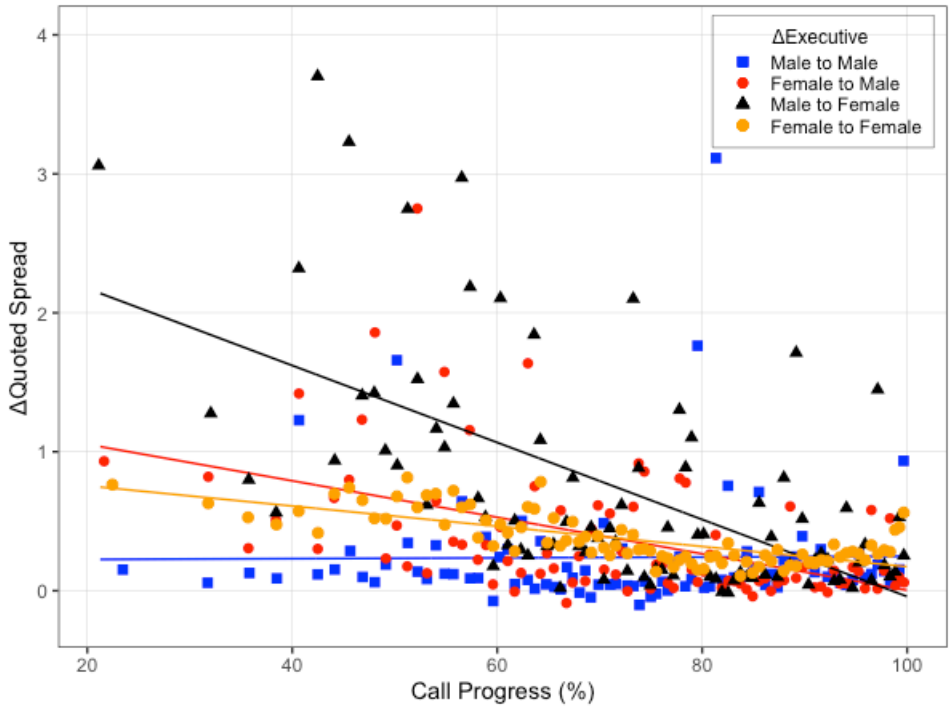
$I(\Delta(\textit{Exec Gender}) = \text{M to F})$ and Call Progress (%). That is, in line with our reasoning above, the later a change from male to female executive speaker occurs in a call, the lower is its impact upon bid-ask spreads. Moreover, risk perceptions decrease throughout an earnings call: Table 5.7 shows that, over the course of an average earnings call, *Quoted Spread* narrows by 38%.

We have checked whether the results in Table 5.7 could be explained by changes in the textual content of responses when the identity of a responding executive changes. Our results appear in Table 5.A.6 in the Appendix. When we include changes in levels of positive, negative, uncertain, and financial tone in our regression specifications, the effect of gender changes reported in Table 5.7 is unchanged.

Second, we investigate the relative impact of the first and subsequent switches from male to female speakers by regressing $\Delta(\textit{Quoted Spread})$ on dummies for the first times gender switches between consecutive executive speakers occur as well as $\Delta \textit{Uptalk}$ and call progress. Table 5.8 reports the results of this regression. Model (1) in the table shows that the average relative change in *Quoted Spread* when the recorded gender of consecutive executive speakers changes from male to female for the first time is 74 percentage points. That compares to the average effect of 37% reported in Table 5.7 for all recorded changes in consecutive executive speakers from male to female. In line with our hypothesis, it follows that after the female executive has spoken for the first time during a call, successive female executive participation has a smaller effect on market risk perceptions. Note that, when we include an interaction term between the effect of the first female speech and the call progress in Model (2), the coefficient on the first female speech term more than doubles. To understand this, note that, with the interaction term, the Model (1) average coefficient on the first female speech is obtained by taking the Model (2) coefficient and adding the interaction term multiplied by the average elapsed call percentage when a female speech first occurred. That calculation implies that the first female speech occurs 56.7% of the way through the call: $0.7403 = 1.693 - 56.7 \times 0.0168$.

Third, we use the Python *binsreg* package to create binned scatterplots of $\Delta(\textit{Quoted Spread})$ against call progress for every change in recorded executive speaker during an earnings call (Cattaneo, Crump, Farrell, and Feng, 2022). That is, for a given change in recorded gender, we take

Figure 5.1. Changes in bid-ask spreads as speaker gender changes through earnings calls
This figure presents binned scatterplots of Δ(*Quoted Spread)* against call progress for every change in recorded executive speaker during an earnings call. The *binsreg* plots include the control variables and fixed effects shown in Table 5.7.



every (Δ(*Quoted Spread)*, call progress (%)) pair in our dataset, divide them into 90 equal-sized bins by call progress, and compute the average Δ(*Quoted Spread)* in each bin. Finally, we regress average Δ(*Quoted Spread)* against Call Progress. Our results are plotted in Figure 5.1.

The plots in Figure 5.1 are consistent with our hypothesis. The black line in Figure 5.1 indicates the relationship between Δ(*Quoted Spread)* and Call Progress when the recorded gender of consecutive executive speakers switches from male to female. The line has a negative slope so that, across our dataset, the later in the call a change from male to female executive speaker occurs, the lower is its impact upon risk perceptions.

5.6. The #MeToo Movement

This section examines the effect of the #MeToo movement upon market reactions to uptalk in male and female speech during earnings calls. The phrase "Me Too" was first used by the activist Tarana Burke to express her empathy for women of colour who had experienced sexual abuse, and

the #MeToo movement achieved global in the wake of a 15 October 2017 social media post by the actress Alyssa Milano in response to a Pulitzer-prize winning New York Times article about the Hollywood producer Harvey Weinstein.[7] The movement highlighted the systemic nature of workplace sexual harassment and of gender discrimination more generally, and it resulted in the resignation of several high-profile male executives who were alleged to have a long history of inappropriate sexual behaviour.

If the #MeToo movement altered gender relationships in the workplace, then it may also have had an effect upon gendered speech patterns. We present summary statistics for pre- and post-#MeToo uptalk levels in Table 5.A.7 in the Appendix. There was a significant decline in uptalk after the start of the #MeToo movement; there was also a substantial drop in the standard deviation of uptalk, and a small decline in median levels of uptalk.

If the #MeToo movement altered attitudes towards gender, then one would expect it to change the way that the gender composition of a call affects gendered speech patterns. We investigate this possibility in Table 5.9, which reports regressions of uptalk amongst female and male executives and analysts on the gender of earnings call participants pre- and post- #MeToo. Neither call composition nor #MeToo has a significant effect upon uptalk levels by analysts (Models (3) and (4)). The first row of Models (1) and (2) shows that an executive's level of uptalk in a earnings call is increasing in the fraction of executives in the call who share his or her gender. This is the effect that we documented in Table : executives appear to be less comfortable when they have the same gender as their peer group. Row three of the table shows that this effect is partly reversed by #MeToo for women, and that it is unaffected for men.

Table 5.9 also reports the effect of female analysts upon uptalk levels. Pre-#MeToo, the proportion of female analysts participating in an earnings call had no effect upon male executive uptalk levels. In contrast, a higher proportion of female analysts participating in the call is associated with a lower level of uptalk by female executives. One possible interpretation of this fact is that, pre-#MeToo, female analysts were viewed by female executives as less threatening than male analysts. Row four of the Table shows that the #MeToo movement had an effect on executives of both

[7]The speed with which the #MeToo movement emerged is demonstrated by Figure 5.A.3 in the Appendix, which reports an explosion in Google searches for #MeToo in the wake of Ms Milano's post.

Table 5.9. The effect of call composition and #MeToo on uptalk levels
This table reports GLM logit estimates of the association between call participant composition and the mean incidence of uptalk among male and female executives and analysts participating in quarterly earnings calls around the #MeToo period, based on the following specification:

$$Mean\ Uptalk_{cftg} = \beta_0 + \beta_1 Call\ Characteristics_{cft} + \beta_2 Call\ Characteristics_{cft} \times MeToo_c + \beta_3 X_{ft} + \lambda_f + \theta_t + \varepsilon_{cft}$$

Here, c and f are indices for the call and the firm, t is the year-quarter of the call, and g is a participant grouping (e.g., all executives, female executives only, male executives only). Standard errors are clustered by year-quarter, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variables:	Mean Uptalk (Female Execs)	Mean Uptalk (Male Execs)	Mean Uptalk (Female Analysts)	Mean Uptalk (Male Analysts)
Model:	(1)	(2)	(3)	(4)
Female Executives (%)	0.6956***	-0.0919***	0.0056	-0.0079
	(0.0843)	(0.0263)	(0.0183)	(0.0177)
Female Analysts (%)	-0.2570***	0.0248	0.0021	-0.0065
	(0.0459)	(0.0151)	(0.0111)	(0.0076)
Female Executives (%) × MeToo	-0.1813**	-0.0544	0.0221	-0.0112
	(0.0855)	(0.0323)	(0.0361)	(0.0379)
Female Analysts (%) × MeToo	0.1854***	0.0943***	-0.0023	0.0251
	(0.0475)	(0.0240)	(0.0261)	(0.0238)
Firm/Stock Controls	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Year-Q FE	✓	✓	✓	✓
Observations	45,974	45,974	45,974	45,974
Adjusted R^2	0.0843	0.1070	0.0975	0.1081

genders: post-#MeToo, analyst gender had a smaller impact on female executive uptalk levels, and male executives actually exhibited higher levels of uptalk when there was a higher proportion of female analysts in the earnings call. We interpret this pair of results as being consistent with a higher level of self confidence amongst female analysts after #MeToo. That self confidence led the analysts to ask tougher questions and so to induce discomfort that raised uptalk levels relative to the pre-#MeToo baseline.

We have also examined the effect of #MeToo upon the relationship between executive uptalk levels and role; our results appear in Table 5.A.8 in the Appendix. The table demonstrates, in line with Table 5.2, that only the CFO role has a consistently positive level of uptalk. We also find that CEOs and board members engage in significantly less uptalk post-#MeToo.

We now examine whether or not the #MeToo movement altered the relationship between executive uptalk and consensus analyst recommendations and earnings forecasts. We do this by regressing consensus average analyst recommendations on average uptalk levels of male and female

Table 5.10. #MeToo Movement, uptalk, and analyst recommendations
This table reports logit estimates of the relationship between uptalk among earnings call participants and subsequent consensus analyst recommendations based on the following specification:

$$Outcome_{cft} = \beta_0 + \beta_1 Mean\ Uptalk_{cftg} + \beta_2 MeToo\ GST_{cft} + \beta_3 Mean\ Uptalk_{cftg} \times MeToo\ GST_{cft} + \beta_4 X_{ft} + \lambda_f + \theta_t + \varepsilon_{cft}.$$

Here, c and f are indices for the call and the firm, t is the year-quarter of the call, and g is a participant grouping (e.g., all executives, female executives only, male executives only). *Outcome* is a dummy variable that captures the consensus analyst recommendation as reported by I/B/E/S. In Models (1) and (2), *Outcome* is 1 in case the consensus estimate is 4 or 5 and, hence, is classified as a buy, and is otherwise 0; in Models (3) and (4), *Outcome* is 1 in case the consensus estimate is 1 or 2 and, hence is classified as a sell, and is otherwise 0. $Mean\ Uptalk_{cftg}$ is the measured level of uptalk for participant group g in the call c. $MeToo\ GST_{cft}$ measures the volume of Google search queries for the term "MeToo" made at the time of the earnings call. X_{ft} is a vector of financial and stock characteristics and consensus analyst recommendations for firm f, and λ_f and θ_t are firm and time fixed effects, respectively. Standard errors are clustered by firm, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variables:	Buy recommendation		Sell recommendation	
	(1)	(2)	(3)	(4)
Mean Uptalk (Female Executives)	-0.3890**	-0.3704**	0.1307**	0.1187**
	(0.1899)	(0.1858)	(0.0586)	(0.0570)
Mean Uptalk (Male Executives)	-0.2042	-0.1710	0.0048	-0.0049
	(0.1561)	(0.1649)	(0.0489)	(0.0482)
MeToo GST	-0.1761	-0.1787	0.1362	0.1362
	(0.4374)	(0.4324)	(0.3193)	(0.3280)
Mean Uptalk (Female Executives) × *MeToo GST*	0.3439	0.3116	0.3894	0.4600
	(0.5953)	(0.5930)	(0.3339)	(0.3443)
Mean Uptalk (Male Executives) × *MeToo GST*	1.273**	1.219**	-0.4323***	-0.4025***
	(0.4987)	(0.5462)	(0.1596)	(0.1553)
Analyst Controls		✓		✓
Firm FE	✓	✓	✓	✓
Year-Q FE	✓	✓	✓	✓
Observations	9,994	9,994	17,088	17,088
Pseudo R^2	0.2603	0.2606	0.2697	0.2702

executives and a Google Search Trends (GST) variable that measures the volume of Google search queries for the term "#MeToo" made around the time of the earnings call.[8] Our results are reported in Table 5.10; they are qualitatively unchanged when we replace the GST variable with a #MeToo dummy that is equal to 1 after October 17 2017 (see Table 5.A.9).

The first row of Table 5.10 reports the effect of pre-#MeToo female executive uptalk upon consensus analyst recommendations. Consistent with Table 5.4, female executive uptalk served significantly to reduce the likelihood of analyst buy recommendations, and to increase the likeli-

[8] Google search trends is a service provided by Google that gives insights into internet search behavior by measuring the volume of internet queries of specific search terms over time and in specific geographic areas. Google aggregates this information for each search term and scales it between 0 to 100, with 0 (100) suggesting lowest (highest) search volume for that search term at a given date and location.

hood of sell recommendations. Once again, in line with Table 5.4, male executive uptalk had no significant association with recommendations before the #MeToo movement. Row four of the table indicates that the effect of female uptalk upon analyst recommendations was not changed by the #MeToo movement. Of course, as per Table 5.A.7, the level of female executive uptalk was significantly lower after #MeToo; but, when that uptalk occurred, it continued to reveal the same concern about earnings as it did before #MeToo, and it was interpreted in the same way.

The fifth row of Table 5.10 reports that, after #MeToo, analysts responded positively to male executive uptalk. That is, while the informational content of female executive uptalk was unchanged by #MeToo, male executive uptalk was received more positively by the market after #MeToo. Our interpretation of this fact relies upon analysis reported by Lins, Roth, Servaes, and Tamayo (2020), who find that firms whose culture does not discriminate on the basis of gender experienced positive excess returns after the #MeToo movement. We conjecture that male executive uptalk is more likely to occur in an organization whose culture discourages sexual discrimination and sexual harassment. If this is the case, then, to the extent that by a firm's officers harms the firm's revenues when it is revealed, male uptalk reveals useful information about future revenue streams. That information has the effect upon analyst recommendations that we report in row five of Table 5.10.[9]

We have demonstrated that analyst recommendations respond positively to male executive uptalk in the post-#MeToo era. Following Section 5.5.1, the results of Table 5.10 could occur for two reasons: they could reflect analysts' gender biases; or they could reflect a rational judgment about future firm earnings. In order to investigate these possibilities, we now investigate the relationship between uptalk and realised earnings. In line with Section 5.5.1, we do so by we regressing SUE and SUEC on executive uptalk; our analysis in this section is supplemented by the inclusion of a #MeToo dummy. Table 5.11 reports our results.

The signs of the coefficients reported in Table 5.11 are the same as those of Table 5.5 but, in contrast to Table 5.5, the SUE coefficients in Table 5.11 are not significant. We contend that this happens because SUE measures unexpected earnings relative to a two-year trend and that the trend

[9]We have also investigated the effect of #MeToo on the relationship between financial constraints, uptalk, and analyst recommendations. Our results appear in Table 5.A.10 in the Appendix; we find that #MeToo has no significant effect in this case.

Table 5.11. #MeToo movement, Uptalk and future unexpected earnings
This table reports OLS estimates of the relationship between unexpected uptalk among earnings call participants and future unexpected earnings before and after the #MeToo movement, based on the following specification:

$$Outcome_{cft} = \beta_0 + \beta_1 \Delta(Mean\ Uptalk)_{cftg} + \beta_2 \Delta(Mean\ Uptalk)_{cftg} \times MeToo_{cft} + \beta_3 X_{ft} + \lambda_f + \theta_t + \varepsilon_{cft}.$$

Here, c and f are indices for the call and the firm, t is the year-quarter of the call, and g is a participant grouping (e.g., all executives, female executives only, male executives only). $Outcome_{ft}$ is a measure of unexpected earnings in the year-quarter after t: in Models (1)–(3), *Outcome* is equal to Standardized Unexpected Earnings (*SUE*) and in Models (4)–(6), *Outcome* is equal to Standardized Unexpected Earnings relative to Consensus analyst estimates (*SUEC*). For firm f in year-quarter t, $SUE_{ft} = (EPS_{f,t} - EPS_{f,t-4} - \mu_{ft}^{SUE})/\sigma_{ft}^{SUE}$, where μ_{ft}^{SUE} and σ_{ft}^{SUE} are the respective mean and standard deviation of $(EPS_{f,\tau} - EPS_{f,\tau-4})$ for $\tau = t-8, t-7, \ldots, t-1$ (Bernard and Thomas, 1990); and $SUEC_{ft} = (EPS_{f,t} - (Consensus\ EPS\ estimate)_{f,t} - \mu_{ft}^{SUEC})/\sigma_{ft}^{SUEC}$, where μ_{ft}^{SUEC}, and σ_{ft}^{SUEC} are the respective mean and standard deviation of $(EPS_{f,\tau} - (Consensus\ EPS\ estimate)_{f,\tau})$ for $\tau = t-8, t-7, \ldots, t-1$. *MeToo* is a dummy variable that is equal to 0 before the #MeToo movement and 1 after the movement starts. X_{ft} is a vector of financial and stock characteristics and consensus analyst recommendations for firm f, and λ_f and θ_t are firm and time fixed effects, respectively. Standard errors are clustered by firm, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variables:	SUE_{t+1}			$SUEC_{t+1}$		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Δ(*Mean Uptalk*) (Female Executives)	-0.0260	-0.0260	-0.0293*	-0.0075**	-0.0076**	-0.0078**
	(0.0175)	(0.0177)	(0.0166)	(0.0034)	(0.0034)	(0.0034)
Δ(*Mean Uptalk*) (Male Executives)	0.0270	0.0322	0.0186	-0.0054	-0.0061	-0.0069
	(0.0266)	(0.0271)	(0.0253)	(0.0049)	(0.0050)	(0.0051)
Δ(*Mean Uptalk*) (Female Executives) × MeToo	-0.0402	-0.0396	-0.0245	0.0078	0.0077	0.0089
	(0.0398)	(0.0398)	(0.0396)	(0.0076)	(0.0076)	(0.0077)
Δ(*Mean Uptalk*) (Male Executives) × MeToo	-0.0347	-0.0362	-0.0190	0.0136*	0.0138*	0.0143*
	(0.0376)	(0.0375)	(0.0354)	(0.0080)	(0.0080)	(0.0081)
Δ(*Mean Uptalk*) (Female Analysts)		0.0305	0.0110		-0.0004	-0.0004
		(0.0358)	(0.0340)		(0.0069)	(0.0070)
Δ(*Mean Uptalk*) (Male Analysts)		-0.0492	-0.0465		0.0069	0.0066
		(0.0350)	(0.0313)		(0.0072)	(0.0073)
Last Recommendation	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Year-Q FE	✓	✓	✓	✓	✓	✓
Firm/Stock controls			✓			✓
Observations	37,944	37,944	37,944	37,970	37,970	37,970
Adjusted R^2	0.045	0.045	0.122	0.026	0.026	0.028
Mean (SD) of Dependent Variable	-0.08 (1.03)			-0.02 (0.34)		

experienced a structural break when the #MeToo movement occurred. To the extent that analysts understood that a structural break had occurred, they factored it into their earnings forecasts. Hence, it seems likely that, by using consensus earnings to normalise unexpected earnings, the SUEC measure corrects to some extent for the structural break and that this is the reason that the SUEC coefficients remain significant in Table 5.11.

Table 5.11 presents new results on the accuracy of analyst responses to male executive uptalk. Recall that, in a regression of SUEC on unexpected executive uptalk across our entire sample, the coefficients for unexpected female executive uptalk are negative and significant. We interpreted

Table 5.12. Uptalk and bid-ask spreads either side of the #MeToo movement
This table presents GLM logit estimates of the relationship between uptalk incidence during executive responses to analyst questions and the mean bid-ask spread of the firm's stock during a given response before and after the start of the #MeToo movement, based on the following regression:

$$\begin{aligned} \textit{Quoted Spread}_{cftpi} = {} & \beta_0 + \beta_1 \textit{Uptalk}_{cftpi} + \beta_2 \textit{Female Executive}_{cftpi} + \beta_3 \textit{Female Analyst}_{cfti} \\ & + \beta_4 \textit{Uptalk}_{cftpi} \times \textit{Female Executive}_{cftpi} + \beta_5 \textit{Uptalk}_{cftpi} \times \textit{Female Analyst}_{cfti} \\ & + \beta_6 \textit{Female Executive}_{cftpi} \times \textit{Female Analyst}_{cfti} \\ & + \beta_7 \textit{Uptalk}_{cftpi} \times \textit{Female Executive}_{cftpi} \times \textit{Female Analyst}_{cfti} + \beta_8 X_{cfti} + \lambda_c + \theta_p + \varepsilon_{cftpi}. \end{aligned}$$

Here, c, f, and p are indices for the call, the firm, and the responding executive, respectively, and i counts question-response pairs from the start of the call. $\textit{Quoted Spread}_{cftpi}$ is the average percentage bid-ask spread for response i by executive p in call c for firm f in year-quarter t; it is calculated using the formula in Equation (5.3). $\textit{Uptalk}_{cftpi}$ is the average level of uptalk during the same response. *Female Analyst* and *Female Executive* are dummies that are equal to 1 in the respective cases where a question is posed by a female analyst and answered by a female executive, and that are otherwise equal to 0. X_{cfti} is a vector, which includes *Call Progress* and *Answer Duration*: $\textit{Call Progress}_{cfti}$ is the percentage of the earnings call that elapsed before response i in call c, expressed as a number between 0 and 100; $\textit{Answer Duration}_{cfti}$ is the duration in minutes of response i in call c. Standard errors are double-clustered by firm and date of the call, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variable:	*Quoted Spread*			
	Pre-#MeToo		Post-#MeToo	
Model:	(1)	(2)	(3)	(4)
Uptalk	-0.0069*	-0.0090**	-0.0037	-0.0030
	(0.0040)	(0.0041)	(0.0025)	(0.0020)
Uptalk × *Female Executive*	0.0293*	0.0307*	0.0313	0.0309
	(0.0159)	(0.0179)	(0.0217)	(0.0260)
Female Analyst	-0.0070	-0.0041	0.0140	0.0043
	(0.0174)	(0.0174)	(0.0174)	(0.0214)
Uptalk × *Female Analyst*		0.0284*		-0.0072
		(0.0140)		(0.0145)
Female Executive × *Female Analyst*		0.0713		0.1197
		(0.0599)		(0.0687)
Uptalk × *Female Executive* × *Female Analyst*		-0.0247		0.0044
		(0.0825)		(0.0665)
Call Progress (%)	-0.0083***	-0.0083***	-0.0052***	-0.0052***
	(0.0011)	(0.0011)	(0.0010)	(0.0010)
Log (*Answer Duration*)	0.0654***	0.0654***	0.0705***	0.0705***
	(0.0056)	(0.0056)	(0.0063)	(0.0063)
Call FE	✓	✓	✓	✓
Executive FE	✓	✓	✓	✓
Observations	249,357	249,357	138,415	138,415
Pseudo R^2	0.762	0.762	0.789	0.789

that result as indicating that analysts consistently fail to change their recommendations as far as they should do in the wake of unexpected female executive uptalk. In line with that result, row one of Models (4)–(5) in Table 5.11 reports a negative pre-#MeToo coefficient for unexpected female executive uptalk, and row three of the table indicates that this effect does not change significantly

after #MeToo. What is new in Table 5.11 are the positive coefficients reported in row four of Models (4)–(5) for post-#MeToo unexpected male executive uptalk. These coefficients, which are significant at the 10% level, indicate that, while analyst recommendations respond positively to unexpected male executive uptalk, analysts fail to update their earnings estimates as far upwards as is warranted. That lack of responsiveness is completely in line with the relationship between female executive uptalk and SUEC.

We now examine the effect of the #MeToo movement on the relationship between executive uptalk and market risk perceptions. Recall from Table 5.6 that, across our entire sample, bid-ask spreads narrow slightly in response to male executive uptalk and widen in response to female executive uptalk. In order to investigate the effects of #MeToo upon these relationships, we split our sample into pre- and post-#MeToo sub-samples, and we run the regression of Table 5.6 on each of those sub-samples. Our results are reported in Table 5.12. The effects reported in Table 5.6 obtain in the pre-#MeToo sample, but they are absent in the post-#MeToo data, where we find no significant relationship between bid-ask spreads and uptalk by executives of either gender. Our explanations for these effects is as follows. First, we have already argued that, because uptalk is a female-typed speech characteristic, post-#MeToo male executive uptalk could arise in organizations in which accusations of sexual harassment are less likely to occur. If that type of organization has a more volatile stock price, then uptalk by men is associated with wider bid-ask spreads after #MeToo. That widening serves to reverse the narrowing identified in the pre-#MeToo data and, hence, could reverse the effect reported in row one of Models (1) and (2) of Table 5.12. Second, in order to explain the absence of a significant effect of female executive uptalk on post-#MeToo bid-ask spreads, recall our argument that the heightened risk perception in Models (1) and (2) was caused by information revelation induced by a lack of confidence. It follows that, if female executive uptalk was less associated with a lack of confidence after #MeToo, then such uptalk ceased to reveal the sort of information that reasonably causes wider bid-ask spreads.

We conclude this section by examining the way that the #MeToo movement affected the within-call dynamics of bid-ask spreads. Following Section 5.5.2, we split our dataset into pre- and post-#MeToo samples, and for each sub-sample we regress $\Delta(\textit{Quoted Spread})$ on dummies for

Table 5.13. The effect of first-time executive gender switches and uptalk on bid-ask spreads around the #MeToo movement

This table reports OLS estimates of the within-earnings-call dynamics of bid-ask spreads on genders of the responding executives, occurrence of uptalk in their responses, and call progress, before and after the start of the #MeToo movement, based on the following regression:

$$\Delta Quoted\ Spread_{cft,n} = \beta_0 + \beta_1 \Delta(Exec\ Gender)_{cft,ns} + \beta_2 \Delta Uptalk_{cft,n} + \beta_3 \Delta(Exec\ Gender)_{cft,ns} \times \Delta Uptalk_{cft,n} + \beta_4 Call\ Progress_{cft,n} + \beta_5 \Delta(Exec\ Gender)_{cft,ns} \times Call\ Progress_{cft,n} + \beta_6 X_{cft} + \lambda_c + \varepsilon_{cft,n}.$$

Here, c and f are indices for the call and the firm, t is the year-quarter of the call, and n counts the transitions between executive speakers in call c for firm f in year-quarter t. As described in the text, $\Delta(Quoted\ Spread)_n$ is the percentage change in the average bid-ask spread for speaker change n in an earnings call. For $s \in \{M\ to\ F, F\ to\ M\}$, $\Delta(Exec\ Gender)_{sn}$ is a dummy that is equal to 1 if gender change s occurs for the first time during the call at speaker change n, and is otherwise equal to 0. $\Delta Uptalk_n$ is the change in uptalk at speaker change n, and $Call\ Progress_n$ is the percentage of the earnings call that has elapsed when change n occurs. $X_{cft,n}$ is a vector of controls that includes $Answer\ Duration_n$ measured in minutes and a $Female\ Analyst_n$ dummy that is equal to 1 if the question in between the speeches at index n is posed by a female analyst and is otherwise 0. Standard errors are double-clustered by firm and date of the call, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variable:	$\Delta(Quoted\ Spread)$	
	Pre #MeToo	Post #MeToo
Model:	(1)	(2)
$I(\Delta(Exec\ Gender) = \text{M to F})_{\text{first}}$	0.3276***	0.4265***
	(0.0990)	(0.1117)
$I(\Delta(Exec\ Gender) = \text{F to M})_{\text{first}}$	0.1217	0.0365
	(0.0765)	(0.0767)
$I(\Delta(Exec\ Gender) = \text{F to F})_{\text{first}}$	-0.0266	-0.0327
	(0.0595)	(0.0759)
$\Delta Uptalk$	0.0162	0.0510***
	(0.0138)	(0.0166)
$I(\Delta(Exec\ Gender) = \text{M to F})_{\text{first}} \times \Delta Uptalk$	-0.0462	0.0482
	(0.0882)	(0.1581)
$I(\Delta(Exec\ Gender) = \text{F to M})_{\text{first}} \times \Delta Uptalk$	-0.0707	0.0194
	(0.0665)	(0.0527)
$I(\Delta(Exec\ Gender) = \text{F to F})_{\text{first}} \times \Delta Uptalk$	-0.0760	0.1458
	(0.0498)	(0.1026)
Call Progress (%)	-0.0034***	-0.0045***
	(0.0008)	(0.0008)
Log (Answer Duration)	-0.0123	0.1354***
	(0.0180)	(0.0208)
Female Analyst	-0.0418	-0.0570
	(0.0382)	(0.0433)
Call FE	✓	✓
Observations	215,526	117,309
Adjusted R^2	0.103	0.145

$\Delta(Mean\ Uptalk)$, on $\Delta Uptalk$, and other variables. Table 5.13 reports our results. In line with Table 5.7, Table 5.13 reports a large and very significant widening in the bid-ask spreads when the gender of the responding executive switches from male to female between responses. The

change is 10 percentage points higher after #MeToo than before it. While Δ*Uptalk* has no effect on Δ(*Quoted Spread)* pre-#MeToo, it has a post-#MeToo effect that is positive and significant at the 1% level. This result is in contrast to the insignificant post-#MeToo effect of *Uptalk* on *Quoted Spread* that is reported in Table 5.12. This contrast arises because the *Uptalk* coefficient in Table 5.12 concerns variation in uptalk during a given executive's response, whereas the Δ*Uptalk* coefficient in Table 5.13 applies only to variation in uptalk that occurs between successive speakers.

We have also used pre- and post-#MeToo sub-samples to perform the regressions in table 5.8 of Δ(*Quoted Spread)* on the first switch in gender between consecutive executive speakers. The results are reported in Table 5.A.12 in the Appendix, and are broadly in line with the findings of Table 5.8.

5.7. Conclusion

This chapter breaks new ground in three ways. First, it examines the price efficiency of market responses to gendered behaviour. Second, it is the first study in which the main explanatory variable is constructed using the tonal variation in audio recordings of speeches. Third, it identifies a new real economic effect of the #MeToo social movement.

We show that the female-typed uptalk speech pattern affects the perception of analysts who observe it in earnings calls and, hence, is reflected in analyst buy and sell recommendations. As far as we are aware, ours is the first study to relate real-world market reactions to gendered behaviour in this way. Other contributions to the literature on gender examine the historic roots of modern gender stereotypes (Galor, Özak, and Sarid, 2020; Tur-Prats, 2019; Alesina, Brioschi, and La Ferrara, 2016; Giuliano, 2015; Alesina, Giuliano, and Nunn, 2013; Santacreu-Vasut, Shoham, and Gay, 2013; Boserup, 1967), the impact of those stereotypes on hiring choices and other judgment-based allocation choices (Coffman, Exley, and Niederle, 2021; Reuben, Sapienza, and Zingales, 2014), and the real effects of female leadership, on the board and elsewhere (Kunze and Miller, 2017; Kim and Starks, 2016; Tate and Yang, 2015; Huang and Kisgen, 2013; Adams and Ferreira, 2009). None of these papers is concerned with market price effects, and, to the extent that they perform causal inference, most use laboratory experiments.

We find that analysts respond negatively when female executives use unexpectedly high levels of uptalk in earnings calls. This response appears to be warranted: realised earnings are lower when women use more uptalk than usual. Our data suggest that, if anything, analyst expectations regarding change less than they should in response to female executive uptalk. We also find that market perceptions of risk increase when female executives speak in earnings calls.

Our results draw on a novel dataset. We perform a tonal analysis of speech during earnings calls, and we record a measure of uptalk along with transcripts of earnings call text and real-time price information. We believe that ours is the first study to use real-world audio data to study the price impact of gendered behaviour.

Our analysis yields a surprising result about the #MeToo social movement. Analyst responses to uptalk by female executives do not change in the wake of #MeToo (although the volume of such uptalk diminishes), but, after #MeToo, analysts respond positively to uptalk by male executives. That is, one of the consequences of a movement that was intended to achieve greater social justice for women was to reward men who exhibited female-typed speech patterns.

We believe that our methods could be used to investigate the real economic effects of cultured mores in other settings. A population can be divided culturally in many ways. When a particular cultural grouping is associated with a way of speaking, a database of potentially market-sensitive speech can be with data about cultural markers and prices to study the real effects of that pattern. This type of analysis has the potential to shed new light on the real economic effects of perceptions not only of gender, but also of race, class, and educational achievements.

5.A. Appendix

Descriptive Statistics

Table 5.A.1a presents dialogue-level descriptive statistics for the 1,607,428 questions and answers in our earnings calls sample. Uptalk occurs more intensely, and varies more, among executives than analysts. The average incidence of uptalk among executives (analysts) is 1.13 (1.02). Moreover, uptalk is more intense, and varies more, among female executives and analysts than among their male counterparts. It is clear from a comparison of the 5th and 95th percentiles of uptalk levels for female executives and analysts that uptalk varies much more among female executives. The distribution of uptalk is more similar for male executives and analysts. The average (median) quoted spread during call Q&As is 0.52% (0.15%). We also report percentage changes in quoted spreads for every consecutive pair of executive responses in a call, grouped by genders of the executive pairs. The highest average spread is 0.28% and occurs when consecutive speakers have different genders.

Table 5.A.1b reports call-level summary statistics. We group call participants by role (executive or analyst) and by gender, and report average uptalk for each grouping. Mean uptalk is marginally higher among executives than among analysts. Mean uptalk is higher by 0.12 units for female than for male executives; there is little difference in uptalk levels for female and male analysts. Female participation in earnings calls as a percentage of all participating executives (analysts) is rather low at just 10% (14%) on average. The average firm in the sample is moderately sized with assets of $12.82 billion, and receives an average consensus recommendation of 2.26 tilting it towards the sell side as confirmed by the large (small) mean value of the *Sell* (*Buy*) recommendation dummy.

Table 5.A.1. Descriptive statistics

This table presents summary statistics for variables listed in the first column. Panel (a) reports statistics of uptalk by participant type and gender and quoted spreads at the dialogue-level for each question and answer in our earnings calls sample. Panel (b) reports call-level statistics of mean uptalk by participant type and gender, financial and stock performance characteristics of firms, and analyst recommendations.

(a) Dialogue-level statistics

Variable	Mean	SD	Min	5p	25p	50p	75p	95p	Max	Obs
Uptalk (*Executive*)	1.13	1.69	0	0.14	0.54	0.83	1.09	2.74	15	1,607,428
Uptalk (*Female Executive*)	1.17	1.76	0	0.13	0.51	0.82	1.11	3.26	15	114,082
Uptalk (*Male Executive*)	1.13	1.69	0	0.14	0.54	0.83	1.09	2.72	15	1,493,346
Uptalk (*Analyst*)	1.02	1.21	0	0.21	0.61	0.86	1.08	1.98	15	1,607,428
Uptalk (*Female Analyst*)	1.04	1.32	0	0.21	0.62	0.86	1.08	2	15	147,582
Uptalk (*Male Analyst*)	1.02	1.20	0	0.21	0.61	0.86	1.08	1.97	15	1,459,846
Quoted Spread (%)	0.52	0.89	0.01	0.02	0.06	0.15	0.45	3.42	3.42	387,930
Δ(*Quoted Spread)* (%): Male to Male	0.21	0.83	-0.94	-0.76	-0.26	0	0.34	2.52	2.52	301,829
Δ(*Quoted Spread)* (%): Male to Female	0.28	0.94	-0.94	-0.80	-0.38	0.01	0.61	2.52	2.52	8,686
Δ(*Quoted Spread)* (%): Female to Male	0.28	0.95	-0.94	-0.80	-0.38	0	0.62	2.52	2.52	8,633
Δ(*Quoted Spread)* (%): Female to Female	0.17	0.74	-0.94	-0.70	-0.13	0	0.13	2.41	2.52	13,825

(b) Call-level statistics

Variable	Mean	SD	Min	5p	25p	50p	75p	95p	Max
Mean Uptalk (Executives)	1.03	0.75	0	0.53	0.73	0.87	1.06	2.01	15
Mean Uptalk (Female Executives)	1.10	1.56	0	0.14	0.57	0.81	1.05	2.68	15
Mean Uptalk (Male Executives)	0.98	0.76	0	0.42	0.65	0.82	1.03	2.03	15
Mean Uptalk (Analysts)	0.99	0.75	0	0.48	0.72	0.87	1.04	1.79	15
Mean Uptalk (Female Analysts)	0.98	0.73	0.02	0.35	0.68	0.88	1.05	1.71	15
Mean Uptalk (Male Analysts)	1.01	0.74	0.01	0.45	0.72	0.89	1.06	1.88	15
Female Executives (%)	0.10	0.19	0	0	0	0	0	0.50	1
Female Analysts (%)	0.14	0.29	0	0	0	0	0	1	1
Assets ($ bil.)	12.82	31.60	0.01	0.10	0.66	2.31	8.32	64.27	190.17
Book/Equity	0.55	0.50	-0.89	0.03	0.23	0.45	0.77	1.42	3.09
Profitability (%)	1.84	4.78	-30.30	-6.56	0.58	2.43	3.96	7.34	12.74
ROA (%)	0.04	5.05	-38.71	-9.11	-0.00	0.84	1.96	4.71	10.69
FinConstraints	2.81	8.17	-32.81	-1.69	0.58	1.79	3.71	12.67	50.58
Momentum(-90,-2)	-0	0.24	-3.54	-0.37	-0.11	-0	0.11	0.35	3.99
Volatility(-90,-2)	0.02	0.01	0	0.01	0.01	0.02	0.03	0.04	0.34
EPS ($)	0.45	1.52	-72.21	-0.67	-0	0.33	0.77	2	74.75
$\Delta\text{EPS}^{t}_{t-4}$ ($)	0.02	0.79	-3.70	-1.03	-0.13	0.03	0.19	1.03	3.31
SUE	-0.05	1.02	-2.47	-1.84	-0.68	-0.04	0.58	1.73	2.48
SUEC	-0.16	1.06	-2.48	-2.04	-0.85	-0.23	0.56	1.62	2.48
Recommendation (mean)	2.26	0.48	1	1.50	1.97	2.25	2.58	3	4.50
Buy Recommendation	0.04	0.21	0	0	0	0	0	0	1
Sell Recommendation	0.90	0.30	0	0	1	1	1	1	1
Observations				52,168					

Table 5.A.2. Example synchronization map for an earnings call

This table shows part of the sychronization map for the Q1 2018 General Motors (NYSE:GM) earnings earnings call; the first four questions by analysts and the corresponding responses by executives appear in the table. The "Speaker Text" column shows the speech fragments (comprising presentations, questions, or answers) spoken by a individual participant during the call. The "Audio Interval" column shows the start and end times of the audio file clip of the Speaker Text.

Question No	Participant Name	ParticipantRole	Speaker Text (truncated for brevity)		Audio Interval (s)
1	John Murphy	Analyst	Just a first question. Now that you're almost through the issues in Korea, just curious what is next on your list to address...	→	[933.60, 963.64]
1	Mary Barra	CEO	First, I would for say we think we have an exceptionally strong franchise in South America. When you look at the market share leadership position...	→	[963.64, 1123.96]
2	John Murphy	Analyst	Got you. That's helpful. And then just a second question around potential for changes in ownership structure, JV requirements in China. I mean, ...	→	[1123.96, 1152.88]
2	Mary Barra	CEO	We think we have an outstanding partner in SAIC. We've been working together for more than 20 years. So we think having a partner that ...	→	[1152.88, 1197.64]
3	John Murphy	Analyst	Okay. And then just lastly real quick on raws. You were sharing – or you were absorbing, I should say, a larger portion of the raw mat complex...	→	[1197.64, 1241.04]
3	Charles Stevens	CFO	Yes, I – John, let me answer that question in a couple of dimensions. First, we buy about $16 billion of raw material on an annual basis...	→	[1241.04, 1359.88]
4	Ryan Brinkman	Analyst	You guys have been very proactive in recent years about exiting under earning or loss-making geographies, ...	→	[1359.88, 1413.32]
4	Mary Barra	CEO	Well, I think already, when you look at – as I mentioned before, in '15 and '16, we launched new very efficient architectures in the mid-sized compact...	→	[1413.32, 1510.36]
4	Charles Stevens	CFO	And if I could just add to that, Ryan, and to Mary's comments. A lot of the questions seem to be focused on the U.S. market, rightfully so...	→	[1510.36, 1579.52]

Table 5.A.3. Factors that influence uptalk among analysts in a call

This table presents OLS estimates of the relationship between the average incidence of uptalk among analysts participating in quarterly earnings calls and call- and firm-level features based on the following specification:

$$Mean\ Uptalk_{cftg} = \beta_0 + \beta_1 FemaleExecs_{cft} + \beta_2 FemaleAnalysts_{cft} + \beta_3 X_{ft} + \lambda_f + \theta_t + \varepsilon_{cft}.$$

Here, c and f are indices for the call and the firm, and t is the year-quarter of the call. $Female\ Executive_{cft}$ and $FemAn_{cft}$ are the respective percentages of female executives and analysts participating in the call. X_{ft} is a vector of financial and stock characteristics and consensus analyst recommendations for firm f, and λ_f and θ_t are firm and time fixed effects, respectively. Standard errors are clustered by firm, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variables:	Mean Uptalk (Analysts)	Mean Uptalk (Female Analysts)		Mean Uptalk (Male Analysts)	
Model:	(1)	(2)	(3)	(4)	(5)
Female Executives(%)	0.0774**	0.0177	0.0160	-0.0026	-0.0017
	(0.0347)	(0.0151)	(0.0152)	(0.0172)	(0.0172)
Female Analysts(%)	-0.0309	-0.0013	-0.0020	-0.0021	-0.0020
	(0.0256)	(0.0083)	(0.0084)	(0.0090)	(0.0095)
Log(Assets)	0.0266*	0.0166***	0.0153***	0.0022	0.0011
	(0.0157)	(0.0042)	(0.0047)	(0.0042)	(0.0042)
Book/Equity	-0.0115	-0.0010	-0.0013	-0.0038	-0.0034
	(0.0137)	(0.0066)	(0.0067)	(0.0032)	(0.0030)
Profitability	-0.1544	-0.0486	-0.1063	-0.0906	-0.0823
	(0.1625)	(0.0816)	(0.0937)	(0.0820)	(0.0839)
ROA	0.1505	-0.0055	0.0428	-0.0227	-0.0451
	(0.1775)	(0.0900)	(0.0784)	(0.0580)	(0.0565)
FinConstraints	-0.0001	-0.0000	-0.0000	-0.0003**	-0.0003**
	(0.0004)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Momentum(-90,-2)	-0.0043	-0.0113	-0.0096	0.0043	0.0040
	(0.0148)	(0.0069)	(0.0073)	(0.0039)	(0.0037)
Volatility(-90,-2)	0.2793	0.5326**	0.6116**	-0.0260	-0.0880
	(0.4054)	(0.2093)	(0.2310)	(0.1414)	(0.1419)
EPS	-0.0025	0.0003	0.0018	0.0005	0.0007
	(0.0024)	(0.0014)	(0.0021)	(0.0007)	(0.0008)
ΔEPS_{t-4}^{t}	0.0007	-0.0000	0.0004	0.0002	0.0001
	(0.0018)	(0.0008)	(0.0008)	(0.0006)	(0.0005)
SUE		0.0011		-0.0006	
		(0.0014)		(0.0010)	
Recent Rec (mean)		0.0005		0.0060	
		(0.0036)		(0.0043)	
SUE_{t+1}			0.0009		-0.0005
			(0.0015)		(0.0015)
Next Rec (mean)			0.0034		0.0025
			(0.0036)		(0.0043)
Firm FE	✓	✓	✓	✓	✓
Year-Q FE	✓	✓	✓	✓	✓
Observations	52,168	45,938	46,366	45,938	46,366
Adjusted R^2	0.0289	0.0976	0.0939	0.1074	0.1092

Table 5.A.4. Speed with which analysts respond to executive uptalk.
This table OLS estimates based on the following specification to investigate the speed with which analyst recommendations respond to unexpected uptalk among earnings call participants:

$$Outcome_{cft} = \beta_0 + \beta_1 \Delta(Mean\ Uptalk)_{cftg} + \beta_2 X_{ft} + \lambda_f + \theta_t + \varepsilon_{cft}.$$

Here, c and f are indices for the call and the firm, t is the year-quarter of the call, and g is a participant grouping (e.g., all executives, female executives only, male executives only). $Outcome_{cft}$ captures the average of individual recommendations issued by analysts within 1–2, 3–4, or 5–6 weeks after the call. $\Delta(Mean\ Uptalk)_{cftg}$ is the unexpected level of uptalk, as defined in Equation (5.2). X_{ft} is a vector of financial and stock characteristics and consensus analyst recommendations for firm f, and λ_f and θ_t are firm and time fixed effects, respectively. Standard errors are clustered by firm, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Post-call period of recommendations:	(Weeks 1–2)	(Weeks 3–4)	(Weeks 5–6)
Model:	(1)	(2)	(3)
$\Delta(Mean\ Uptalk)$ (Female Executives)	-0.0461***	-0.0420	-0.0164
	(0.0152)	(0.0423)	(0.0165)
$\Delta(Mean\ Uptalk)$ (Male Executives)	0.0827	0.0250	-0.0312
	(0.0493)	(0.0378)	(0.0283)
Call/Firm/Stock Controls	✓	✓	✓
Analyst Controls	✓	✓	✓
Recent Recommendation	✓	✓	✓
Firm FE	✓	✓	✓
Year-Q FE	✓	✓	✓
Observations	7,315	7,905	8,705
Adjusted R^2	0.096	0.099	0.077

Table 5.A.5. Uptalk and analyst recommendation revisions

This table reports logit estimates of the relationship between unexpected uptalk among earnings call participants and subsequent analyst recommendation changes based on the following specification:

$$Outcome_{cft} = \beta_0 + \beta_1 \Delta(Mean\ Uptalk)_{cftg} + \beta_2 X_{ft} + \lambda_f + \theta_t + \varepsilon_{cft}.$$

Here, *c* and *f* are indices for the call and the firm, *t* is the year-quarter of the call, and *g* is a participant grouping (e.g., all executives, female executives only, male executives only). $Outcome_{cft}$ is a dummy variable that captures major changes in the consensus analyst recommendation as reported by I/B/E/S. In Models (1)-(3), *Outcome* is 1 in case the consensus estimate is revised from 1 or 2 before the call to 4 or 5 afterwards and, hence, is classified as an upgrade to buy, and is otherwise 0; in Models (4)-(6), *Outcome* is 1 in case the consensus estimate is revised from 4 or 5 before the call to 1 or 2 afterwards and, hence, is classified as a downgrade to sell, and is otherwise 0. $\Delta(Mean\ Uptalk)_{cftg}$ is the unexpected level of uptalk, as defined in Equation (5.2). X_{ft} is a vector of financial and stock characteristics and consensus analyst recommendations for firm *f*, and λ_f and θ_t are firm and time fixed effects, respectively. Standard errors are clustered by firm, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variables:	Recommendation change (upgrade to buy)			Recommendation change (downgrade to sell)		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta(Mean\ Uptalk)$ (Female Executives)	-0.2477	-0.2508	-0.3749	0.3823***	0.3851***	0.4160***
	(0.3453)	(0.3529)	(0.5177)	(0.1400)	(0.1396)	(0.1071)
$\Delta(Mean\ Uptalk)$ (Male Executives)	0.0414	0.0321	-0.3561	-0.6025	-0.4368	-0.8242
	(0.3683)	(0.4026)	(0.6172)	(0.6943)	(0.6826)	(0.9881)
$\Delta(Mean\ Uptalk)$ (Female Analysts)		-0.0683	0.0703		-0.6236	-0.3036
		(0.2887)	(0.2476)		(0.5992)	(0.5075)
$\Delta(Mean\ Uptalk)$ (Male Analysts)		0.0657	0.1540		-0.8443	-0.8831
		(0.5446)	(0.7077)		(0.8385)	(1.023)
Firm/Stock controls			✓			✓
Last Recommendation	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Year-Q FE	✓	✓	✓	✓	✓	✓
Observations	3,446	3,446	3,352	2,762	2,762	2,659
Pseudo R^2	0.0581	0.0582	0.0842	0.0695	0.0723	0.3919

Table 5.A.6. Effect of changes in speaker gender and textual tone on bid-ask spreads

This table reports OLS estimates of the within-earnings-call dynamics of bid-ask spreads on genders of the responding executives, occurrence of uptalk in their responses, and call progress, based on the following regression:

$$\Delta(\textit{Quoted Spread})_{cft,n} = \beta_0 + \beta_1 \Delta(\textit{Exec Gender})_{cft,n} + \beta_2 \Delta(\textit{Tone})_{cft,n} + \beta_3 \Delta(\textit{Exec Gender})_{cft,n} \times \Delta(\textit{Tone})_{cft,n} + \beta_4 \textit{CallProgress}_{cft,n} + \beta_5 \Delta(\textit{Exec Gender})_{cft,n} \times \textit{Call Progress}_{cft,n} + \beta_6 X_{cft,n} + \lambda_c + \varepsilon_{cft}.$$

Here, c and f are indices for the call and the firm, t is the year-quarter of the call, and n counts the transitions between executive speakers in call c for firm f in year-quarter t. As described in the text, $\Delta(\textit{Quoted Spread})_n$ is the percentage change in the average bid-ask spread for speaker change n in an earnings call, $\Delta \textit{Uptalk}_n$ is the change in uptalk, and $\textit{Call Progress}_n$ is the percentage of the earnings call that has elapsed when speaker change n occurs. $\Delta(\textit{Tone})_n \in \{\textit{Positive}, \textit{Negative}, \textit{Uncertain}, \textit{Financial}\}$ is the change in textual tone that occurs when speaker change n occurs. $X_{cft,n}$ is a vector of controls that includes $\textit{Answer Duration}_n$ measured in minutes and a $\textit{Female Analyst}_n$ dummy that is equal to 1 if the question in between the speeches at index n is posed by a female analyst and is otherwise 0. Standard errors are double-clustered by firm and date of the call, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variable:	ΔQuoted Spread$_{n-1,n}$			
Model:	(1)	(2)	(3)	(4)
ΔExecutive(Male$_{n-1}$ to Female$_n$)	0.3517***	0.4177***	0.3319***	0.4023***
	(0.0857)	(0.0854)	(0.0927)	(0.0845)
ΔExecutive(Female$_{n-1}$ to Male$_n$)	0.0639	0.1349**	0.0827	0.0829
	(0.0585)	(0.0576)	(0.0660)	(0.0601)
ΔExecutive(Female$_{n-1}$ to Female$_n$)	-0.0241	0.0269	-0.0256	0.0210
	(0.0578)	(0.0520)	(0.0643)	(0.0530)
ΔPositive Tone	0.0018			
	(0.0049)			
ΔExecutive(Male$_{n-1}$ to Female$_n$) $\times$ ΔPos Tone	-0.0602			
	(0.0513)			
ΔExecutive(Female$_{n-1}$ to Male$_n$) $\times$ ΔPos Tone	-0.0151			
	(0.0151)			
ΔExecutive(Female$_{n-1}$ to Female$_n$) $\times$ ΔPos Tone	0.0040			
	(0.0145)			
ΔNegative Tone		0.0050		
		(0.0037)		
ΔExecutive(Male$_{n-1}$ to Female$_n$) $\times$ ΔNeg Tone		0.0021		
		(0.0159)		
ΔExecutive(Female$_{n-1}$ to Male$_n$) $\times$ ΔNeg Tone		-0.0015		
		(0.0093)		
ftof $\times$ ΔNeg Tone		0.0016		
		(0.0102)		
ΔUncertain Tone			-0.0046	
			(0.0072)	
ΔExecutive(Male$_{n-1}$ to Female$_n$) $\times$ ΔUnc Tone			-0.0045	
			(0.0687)	
ΔExecutive(Female$_{n-1}$ to Male$_n$) $\times$ ΔUnc Tone			0.0067	
			(0.0174)	
ΔExecutive(Female$_{n-1}$ to Female$_n$) $\times$ ΔUnc Tone			-0.0149	
			(0.0220)	
ΔFinancial Tone				-0.0051
				(0.0053)
ΔExecutive(Male$_{n-1}$ to Female$_n$) $\times$ ΔFin Tone				0.0354
				(0.0424)
ΔExecutive(Female$_{n-1}$ to Male$_n$) $\times$ ΔFin Tone				0.0029
				(0.0141)
ΔExecutive(Female$_{n-1}$ to Female$_n$) $\times$ ΔFin Tone				-0.0316*
				(0.0179)
Call Progress (%)	-0.0032***	-0.0030***	-0.0022***	-0.0030***
	(0.0006)	(0.0006)	(0.0007)	(0.0006)
Log (Answer Duration)	0.0753***	0.0797***	0.0742***	0.0791***
	(0.0129)	(0.0139)	(0.0154)	(0.0136)
Female Analyst	-0.0670**	-0.0658**	-0.0447	-0.0485
	(0.0318)	(0.0311)	(0.0372)	(0.0336)
Call FE	✓	✓	✓	✓
Observations	254,862	275,155	206,862	254,189
Adjusted R^2	0.124	0.120	0.125	0.130

Table 5.A.7. Effect of #MeToo on uptalk levels by role and gender

This table presents descriptive statistics of the incidence of uptalk in for speech fragments uttered by make and female executives and analysts in the pre- and post-#MeToo periods.

	Pre #MeToo				Post #MeToo				
Variable	Obs	Mean	SD	Median	Obs	Mean	SD	Median	*t*-test Diff in means
Male analyst	397894	1.186	1.786	0.865	190040	0.993	1.099	0.856	50.82***
Female analyst	43641	1.283	2.035	0.876	18698	0.993	1.015	0.871	23.68***
Male executive	958320	1.147	1.688	0.851	558199	1.003	1.193	0.817	60.99***
Female executive	76712	1.188	1.683	0.872	50348	1.014	1.243	0.833	21.13***

Table 5.A.8. Uptalk by participant type and role: the impact of the #MeToo movement

Table reports OLS estimates of call-level features and executive roles associated with the incidence of uptalk in each speech fragment (presentations, questions, and answers) of quarterly earnings calls based on the following specification:

$$\begin{aligned} Uptalk_{cpi} = \beta_0 &+ \beta_1 Female_{cp} + \beta_2 Executive_{cp} + \beta_3 Executive_{cp} \times Female_{cp} \beta_4 Female_{cp} \times MeToo_c \\ &+ \beta_5 Executive_{cp} \times MeToo_c + \beta_6 Executive_{cp} \times Female_{cp} \times MeToo_c + \beta_7 X_{ci} \\ &+ \beta_8 X_{ci} \times Female_{cp} + \beta_9 X_{ci} \times MeToo_c + \beta_{10} X_{ci} \times Female_{cp} \times MeToo_c + \lambda_c + \varepsilon_{cpi} \end{aligned}$$

Here, c and p are indices for the call and for each individual in the call, and i is a count of speech fragments starting from the beginning of the call. *Female* is a dummy equal to 1 if the call participant is female. *Executive* is a dummy equal to 1 if the call participant is the firm's executive, and 0 if the participant is an analyst. Q&A is a dummy equal to 1 if the speech fragment is a question or answer, and 0 if the speech fragment is a presentation made by the firm's executive. X_{ci} and λ_c denote call-specific controls and call fixed effects, respectively.

Dependent Variable:	Uptalk			
	(1)	(2)	(3)	(4)
Female	0.011***	0.008***	0.009***	0.009***
	(0.002)	(0.003)	(0.001)	(0.001)
Executive	0.325***	0.315***	0.324***	0.324***
	(0.008)	(0.007)	(0.008)	(0.008)
Executive × Female	0.009	0.016	0.014	0.014
	(0.020)	(0.015)	(0.020)	(0.020)
Female × MeToo	-0.183***	-0.182***	-0.181***	-0.181***
	(0.007)	(0.009)	(0.007)	(0.007)
Executive × Female × MeToo	0.196***	0.199***	0.191***	0.191***
	(0.027)	(0.020)	(0.027)	(0.027)
CEO	0.021*	0.006	0.019	0.019
	(0.012)	(0.008)	(0.012)	(0.012)
CEO × Female	0.011		0.005	0.005
	(0.050)		(0.051)	(0.051)
CEO × MeToo	-0.040***		-0.035**	-0.035**
	(0.013)		(0.015)	(0.015)
CEO × Female × MeToo	-0.028		-0.031	-0.031
	(0.053)		(0.050)	(0.050)
CFO	0.029**	0.034***	0.030**	0.030**
	(0.013)	(0.008)	(0.013)	(0.013)
CFO × Female	0.013		0.012	0.012
	(0.036)		(0.037)	(0.037)
CFO × MeToo	0.008		0.006	0.006
	(0.015)		(0.015)	(0.015)
CFO × Female × MeToo	0.021		0.025	0.025
	(0.044)		(0.044)	(0.044)
Chair	0.004	0.020	0.006	0.006
	(0.012)	(0.016)	(0.017)	(0.017)
Chair × Female		0.024	0.021	0.021
		(0.097)	(0.105)	(0.105)
Chair × MeToo		-0.041**	-0.006	-0.006
		(0.017)	(0.020)	(0.020)
Chair × Female × MeToo		-0.048	-0.007	-0.007
		(0.107)	(0.112)	(0.112)
Board	0.022	0.127	0.128	0.128
	(0.038)	(0.081)	(0.081)	(0.081)
Board × Female		-0.071	-0.069	-0.069
		(0.349)	(0.348)	(0.348)
Board × MeToo		-0.172*	-0.175**	-0.175**
		(0.089)	(0.089)	(0.089)
Board × Female × MeToo		0.296	0.284	0.284
		(0.349)	(0.349)	(0.349)
Dialogue-level controls	✓	✓	✓	✓
Call FE	✓	✓	✓	✓
Observations	1,607,428	1,607,428	1,607,428	1,607,428
Adjusted R^2	0.267	0.267	0.267	0.267

Table 5.A.9. #MeToo Movement, uptalk, and analyst recommendations
This table reports logit estimates of the relationship between uptalk among earnings call participants and subsequent analyst recommendations based on the following specification:

$$Outcome_{cft} = \beta_0 + \beta_1 Mean\ Uptalk_{cftg} + \beta_2 MeToo_{cft} + \beta_3 Mean\ Uptalk_{cftg} \times MeToo_{cft} + \beta_4 X_{ft} + \lambda_f + \theta_t + \varepsilon_{cft}.$$

Here, *c* and *f* are indices for the call and the firm, *t* is the year-quarter of the call, and *g* is a participant grouping (e.g., all executives, female executives only, male executives only). *Outcome* is a dummy variable that captures the consensus analyst recommendation as reported by I/B/E/S. In Models (1) and (2), *Outcome* is 1 in case the consensus estimate is 4 or 5 and, hence, is classified as a buy, and is otherwise 0; in Models (3) and (4), *Outcome* is 1 in case the consensus estimate is 1 or 2 and, hence is classified as a sell, and is otherwise 0. $Mean\ Uptalk_{cftg}$ is the measured level of uptalk for participant group *g* in the call *c*. $MeToo_{cft}$ is a dummy that is equal to 1 if call *c* for firm *f* in year-quarter *t* is on or after the October 15, 2017 start of the #MeToo movement. X_{ft} is a vector of financial and stock characteristics and consensus analyst recommendations for firm *f*, and λ_f and θ_t are firm and time fixed effects, respectively. Standard errors are clustered by firm, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variables:	Buy recommendation		Sell recommendation	
	(1)	(2)	(3)	(4)
Mean Uptalk (Female Executives)	-0.3953*	-0.3763*	0.1418**	0.1268**
	(0.2248)	(0.2231)	(0.0651)	(0.0618)
Mean Uptalk (Male Executives)	-0.3587	-0.3601	0.0229	0.0071
	(0.2496)	(0.2512)	(0.0602)	(0.0601)
MeToo	0.4152*	0.4294*	0.3243**	0.3194**
	(0.2242)	(0.2366)	(0.1342)	(0.1357)
Mean Uptalk (Female Executives) × *MeToo*	0.1239	0.1195	0.0765	0.1008
	(0.2627)	(0.2674)	(0.0976)	(0.0984)
Mean Uptalk (Female Executives) × *MeToo*	0.5096*	0.5385**	-0.1414*	-0.1238*
	(0.2610)	(0.2663)	(0.0722)	(0.0702)
Analyst Controls		✓		✓
Firm FE	✓	✓	✓	✓
Year-Q FE	✓	✓	✓	✓
Observations	9,994	9,994	17,088	17,088
Pseudo R^2	0.2604	0.2608	0.2697	0.2701

Table 5.A.10. #MeToo movement, uptalk and analyst recommendations: the role of firm financial constraints

This table reports logit estimates of the relationship between uptalk among earnings call participants, firm financial constraints, and subsequent analyst recommendations based on the following specification:

$$Outcome_{cft} = \beta_0 + \beta_1 Mean\ Uptalk_{cftg} + \beta_2 MeToo + \beta_3 Financial\ Constraints_{ft} + \beta_4 Mean\ Uptalk_{cftg} \times Financial\ Constraints_{ft} \times MeToo + \beta_5 X_{ft} + \lambda_f + \theta_t + \varepsilon_{cft}.$$

Here, *c* and *f* are indices for the call and the firm, *t* is the year-quarter of the call, and *g* is a participant grouping (e.g., all executives, female executives only, male executives only). $Outcome_{cft}$ is a dummy variable that captures the consensus analyst recommendation as reported by I/B/E/S. In Models (1) and (2), *Outcome* is 1 in case the consensus estimate is 4 or 5 and, hence, is classified as a buy, and is otherwise 0; in Models (3) and (4), *Outcome* is 1 in case the consensus estimate is 1 or 2 and, hence is classified as a sell, and is otherwise 0. $Mean\ Uptalk_{cftg}$ is the measured level of uptalk for participant group *g* in the call *c*. $Financial\ Constraints_{ft}$ denotes financial constraints facing the firm *f* at time *t*, which are measured using the method developed by Lamont, Polk, and Saaá-Requejo (2001) as described in Section 5.3.1. *MeToo* is a dummy equal to 1 if the call occurred on or after October 15, 2017 when the *MeToo* movement became a widespread social phenomenon. X_{ft} is a vector of financial and stock characteristics and consensus analyst recommendations for firm *f*, and λ_f and θ_t are firm and time fixed effects, respectively. Standard errors are clustered by firm, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variables:	Buy recommendation		Sell recommendation	
Model:	(1)	(2)	(3)	(4)
Δ(*Mean Uptalk*) (Female Executives)	-0.1990	-0.2460	0.2536	0.3368
	(0.2508)	(0.3267)	(0.2346)	(0.2680)
Δ(*Mean Uptalk*) (Male Executives)	-0.1936	-0.2072	0.1520	0.1784
	(0.1816)	(0.1954)	(0.1091)	(0.1123)
MeToo	0.1854		-0.1583	
	(0.1429)		(0.1036)	
FinConstraints × MeToo	-0.0034	-0.0002	0.0153	0.0118
	(0.0142)	(0.0149)	(0.0103)	(0.0101)
Δ(*Mean Uptalk*) (Female Executives) × FinConstraints	-0.0554	-0.0527	0.0178	0.0115
	(0.0449)	(0.0592)	(0.0300)	(0.0318)
Δ(*Mean Uptalk*) (Female Executives) × MeToo	0.1481	0.2493	-0.4414	-0.5307
	(0.5309)	(0.5654)	(0.3950)	(0.4157)
Δ(*Mean Uptalk*) (Female Executives) × FinConstraints × MeToo	-0.1018	-0.1113	0.0505	0.0603
	(0.0707)	(0.0818)	(0.0657)	(0.0660)
Δ(*Mean Uptalk*) (Male Executives) × FinConstraints ×	-0.0251	-0.0295	-0.0116	-0.0015
	(0.0304)	(0.0295)	(0.0241)	(0.0237)
Δ(*Mean Uptalk*) (Male Executives) × MeToo	0.0785	0.0973	-0.4191**	-0.4289**
	(0.2249)	(0.2370)	(0.1972)	(0.1973)
Δ(*Mean Uptalk*) (Male Executives) × FinConstraints × MeToo	0.0565	0.0580*	0.0170	0.0085
	(0.0351)	(0.0349)	(0.0264)	(0.0261)
Firm/Stock Controls	✓	✓	✓	✓
Mean Uptalk (Analysts)	✓	✓	✓	✓
Recent Recommendation	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Year-Q FE		✓		✓
Observations	9,969	9,968	17,012	17,008
Pseudo R^2	0.671	0.679	0.672	0.678

Table 5.A.11. Uptalk and bid-ask spreads by executive role

This table presents GLM logit estimates of the relationship between uptalk incidence during executive responses by role to analyst questions and the mean bid-ask spread of the firm's stock during a given response, based on the following regression:

$$\begin{aligned} \textit{Quoted Spread}_{cftpi} = {} & \beta_0 + \beta_1 \textit{Uptalk}_{cftpi} + \beta_2 \textit{Female Executive}_{cftpi} + \beta_3 \textit{Female Analyst}_{cfti} \\ & + \beta_4 \textit{Uptalk}_{cftpi} \times \textit{Female Executive}_{cftpi} + \beta_5 \textit{Uptalk}_{cftpi} \times \textit{Female Analyst}_{cfti} \\ & + \beta_6 \textit{Female Executive}_{cftpi} \times \textit{Female Analyst}_{cfti} \\ & + \beta_7 \textit{Uptalk}_{cftpi} \times \textit{Female Executive}_{cftpi} \times \textit{Female Analyst}_{cfti} + \beta_8 X_{cfti} + \lambda_c + \theta_p + \varepsilon_{cftpi}. \end{aligned}$$

Here, c, f, and p are indices for the call, the firm, and the responding executive, respectively, and i counts question-response pairs from the start of the call. $\textit{Quoted Spread}_{cftpi}$ is the average percentage bid-ask spread for response i by executive p in call c for firm f in year-quarter t; it is calculated using the formula in Equation (5.3). $\textit{Uptalk}_{cftpi}$ is the average level of uptalk during the same response. *Female Analyst* and *Female Executive* are dummies that are equal to 1 in the respective cases where a question is posed by a female analyst and answered by a female executive, and that are otherwise equal to 0. X_{cfti} is a vector, which includes *Call Progress* and *Answer Duration*: $\textit{Call Progress}_{cfti}$ is the percentage of the earnings call that elapsed before response i in call c, expressed as a number between 0 and 100; $\textit{Answer Duration}_{cfti}$ is the duration in minutes of response i in call c. Standard errors are double-clustered by firm and date of the call, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variable:	*Quoted Spread*			
Responding Executives	CEOs	CFOs	Chairpersons	Board Members
Model:	(1)	(2)	(3)	(4)
Uptalk (Executive)	-0.0072*	-0.0099	-0.0138*	0.0045
	(0.0043)	(0.0071)	(0.0073)	(0.0223)
Uptalk (Executive) × Female Executive	0.0491**	0.0104	0.0646	-0.0729
	(0.0238)	(0.0219)	(0.0478)	(0.2196)
Female Analyst	0.0020	-0.0055	-0.0214	0.0443
	(0.0140)	(0.0218)	(0.0242)	(0.0962)
Uptalk (Executive) × Female Analyst	0.0181	0.0069	0.0253	-0.0022
	(0.0114)	(0.0194)	(0.0183)	(0.0971)
Female Executive × Female Analyst	0.0462	0.0378	0.0306	-0.0462
	(0.0679)	(0.0712)	(0.1148)	(0.1020)
Uptalk (Executive) × Female Executive × Female Analyst	-0.0481	-0.0408	-0.0490	0.0196
	(0.0818)	(0.0892)	(0.1393)	(0.1467)
Call Progress (%)	-0.0062***	-0.0080***	-0.0085***	-0.0084***
	(0.0004)	(0.0005)	(0.0006)	(0.0019)
Log (Answer Duration)	0.0659***	0.0628***	0.0622***	0.0215
	(0.0050)	(0.0080)	(0.0072)	(0.0237)
Call FE	✓	✓	✓	✓
Executive FE	✓	✓	✓	✓
Observations	222,006	103,817	85,487	6,110
Pseudo R^2	0.759	0.802	0.770	0.783

Table 5.A.12. The effect of first-time executive gender switches and uptalk on bid-ask spreads around the #MeToo movement

This table reports OLS estimates of the within-earnings-call dynamics of bid-ask spreads on genders of the responding executives, occurrence of uptalk in their responses, and call progress, before and after the start of the #MeToo movement based on the following regression:

$$\Delta Quoted\ Spread_{cft,n} = \beta_0 + \beta_1 \Delta(Exec\ Gender)_{cft,ns} + \beta_2 \Delta Uptalk_{cft,n} + \beta_3 \Delta(Exec\ Gender)_{cft,ns} \times \Delta Uptalk_{cft,n} + \beta_4 Call\ Progress_{cft,n} + \beta_5 \Delta(Exec\ Gender)_{cft,ns} \times Call\ Progress_{cft,n} + \beta_6 X_{cft} + \lambda_c + \varepsilon_{cft,n}.$$

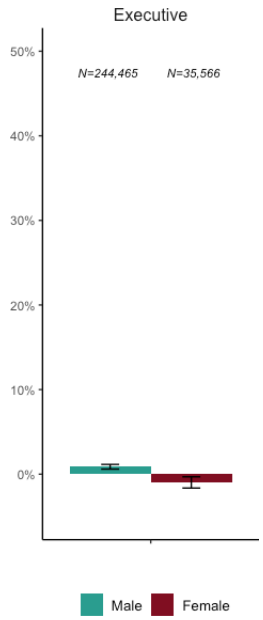
Here, c and f are indices for the call and the firm, t is the year-quarter of the call, and n counts the transitions between executive speakers in call c for firm f in year-quarter t. As described in the text, $\Delta(Quoted\ Spread)_n$ is the percentage change in the average bid-ask spread for speaker change n in an earnings call. For $s \in \{M\ to\ F, F\ to\ M\}$, $\Delta(Exec\ Gender)_{sn}$ is a dummy that is equal to 1 if gender change s occurs for the first time during the call at speaker change n, and is otherwise equal to 0. $\Delta Uptalk_n$ is the change in uptalk at speaker change n, and $Call\ Progress_n$ is the percentage of the earnings call that has elapsed when change n occurs. $X_{cft,n}$ is a vector of controls that includes $Answer\ Duration_n$ measured in minutes and a $Female\ Analyst_n$ dummy that is equal to 1 if the question in between the speeches at index n is posed by a female analyst and is otherwise 0. Standard errors are double-clustered by firm and date of the call, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Dependent Variable:	$\Delta(Quoted\ Spread)$			
	Pre-#MeToo	Post-#MeToo	Pre-#MeToo	Post-#MeToo
Model:	(1)	(2)	(3)	(4)
$I(\Delta(Exec\ Gender) = \text{M to F})_{\text{first}}$	1.938***	1.323*		
	(0.6658)	(0.7039)		
$I(\Delta(Exec\ Gender) = \text{F to M})_{\text{first}}$			1.567***	-0.1954
			(0.5305)	(0.4342)
Call Progress (%)	-0.0029***	-0.0039***	-0.0031***	-0.0045***
	(0.0008)	(0.0008)	(0.0008)	(0.0008)
$\Delta Uptalk$	0.0107	0.0499***	0.0132	0.0578***
	(0.0130)	(0.0155)	(0.0131)	(0.0165)
$I(\Delta(Exec\ Gender) = \text{M to F})_{\text{first}} \times$ Call Progress (%)	-0.0212**	-0.0103		
	(0.0089)	(0.0107)		
$I(\Delta(Exec\ Gender) = \text{M to F})_{\text{first}} \times \Delta Uptalk$	-0.0344	0.6253		
	(0.2266)	(0.4175)		
$I(\Delta(Exec\ Gender) = \text{F to M})_{\text{first}} \times$ Call Progress (%)			-0.0194***	0.0030
			(0.0063)	(0.0080)
$I(\Delta(Exec\ Gender) = \text{F to M})_{\text{first}} \times \Delta Uptalk$			-0.2469	0.1127
			(0.1581)	(0.1507)
Log (Answer Duration)	-0.0114	0.1364***	-0.0101	0.1389***
	(0.0179)	(0.0206)	(0.0179)	(0.0206)
Female Analyst	-0.0418	-0.0547	-0.0428	-0.0541
	(0.0381)	(0.0432)	(0.0381)	(0.0432)
Call FE	✓	✓	✓	✓
Observations	215,712	118,278	215,712	118,278
Adjusted R^2	0.103	0.145	0.103	0.144

Figure 5.A.1. Incidence of uptalk by gender and participant type

Average levels of uptalk are presented for executive and analyst speakers by gender; 95% confidence interval for standard errors are indicated by whiskers. Each call comprises a presentation, followed by a Q&A session; analysts do not speak during the presentation.

(a) Uptalk during presentations by gender



(b) Uptalk during Q&A by gender

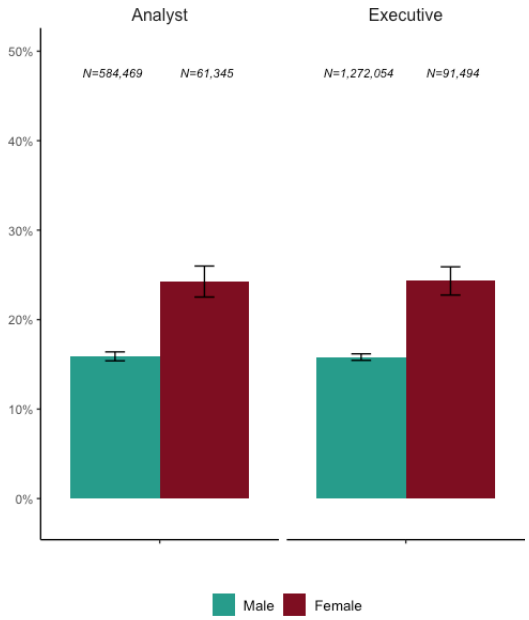
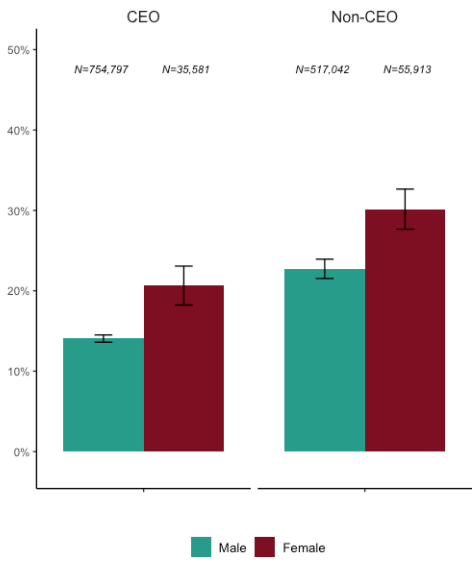


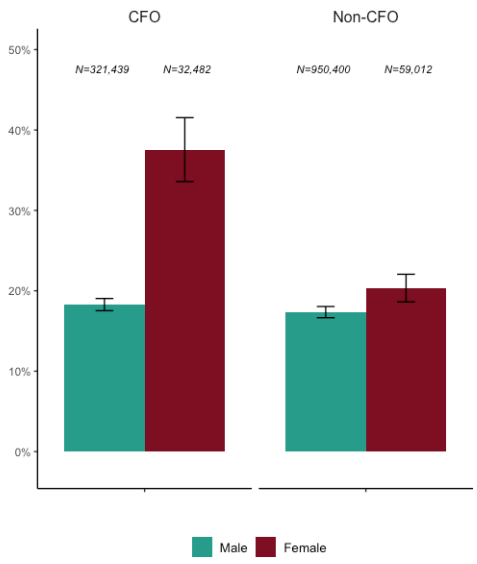
Figure 5.A.2. Incidence of uptalk among executives by gender and role

Average levels of uptalk are presented for executive and analyst speakers by gender; 95% confidence interval for standard errors are indicated by whiskers. Each call comprises a presentation, followed by a Q&A session; analysts do not speak during the presentation.

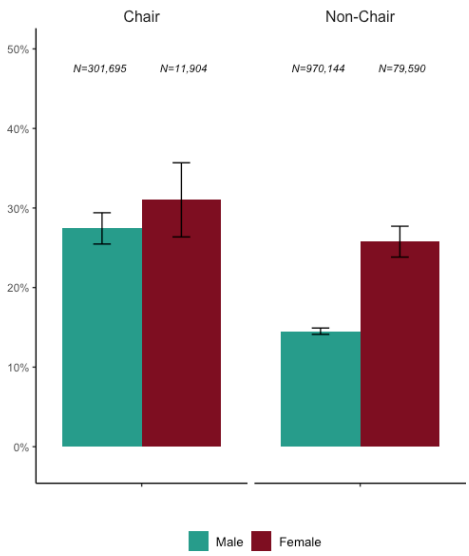
(a) Uptalk among CEOs and non-CEOs



(b) Uptalk among CFOs and non-CFOs



(c) Uptalk among Chair- and non-chairpersons



(d) Uptalk among Board- and non-board members

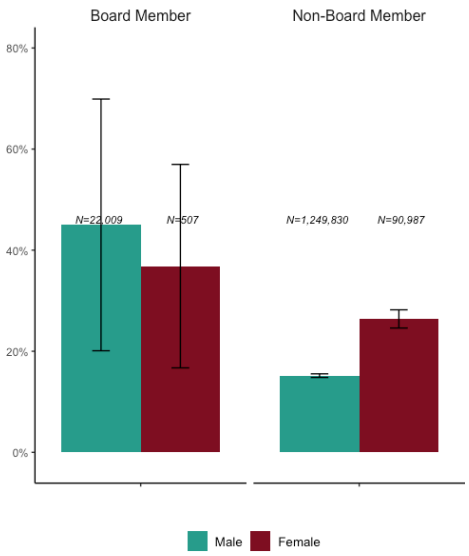


Figure 5.A.3. Google search trends for #MeToo

This Figure shows daily Google search trends in the USA for the search term "metoo" during the sample period. Google collects and aggregates this information based on search behaviour among internet users and scales it between 0 and 100. The lowest search volume for a search term at a given date and location is indicated by a search trend value of 0, and the highest volume by a value of 100.

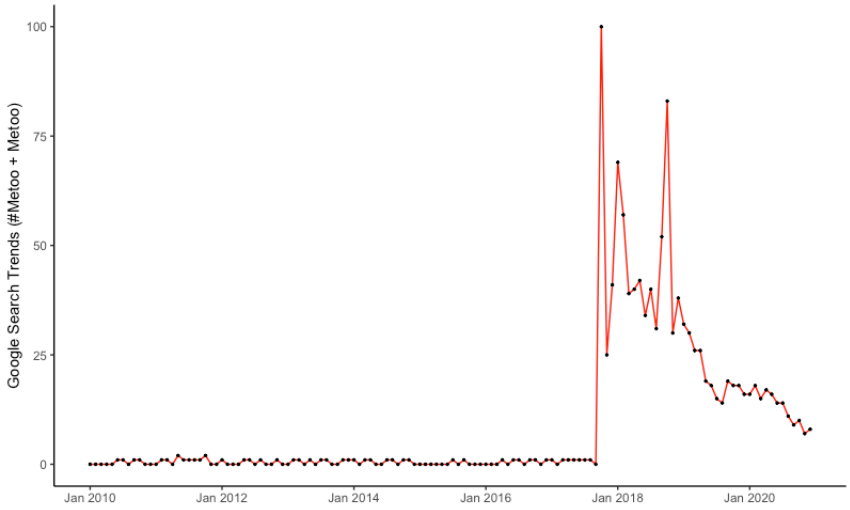


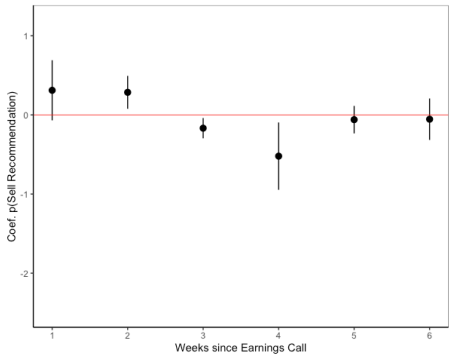
Figure 5.A.4. Uptalk and analyst recommendations by week

This table reports OLS estimates of the relationship between unexpected uptalk among earnings call participants and subsequent analyst recommendations based on the following specification:

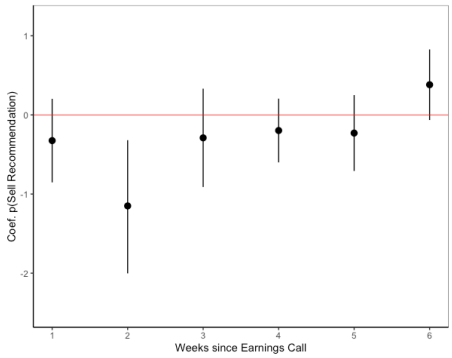
$$Outcome_{cftw} = \beta_0 + \beta_1 \Delta(Mean\ Uptalk)_{cftg} + \beta_2 X_{ft} + \lambda_f + \theta_t + \varepsilon_{cft}.$$

Here, c and f are indices for the call and the firm, t is the year-quarter of the call, g is a participant grouping (e.g., all executives, female executives only, male executives only), and w is a count of the number of weeks after call c. $Outcome_{cftw}$ is the average of individual analyst recommendations for firm f in week w: it is coded 1 for average buy recommendations (I/B/E/S score 4 or 5) in figures (a) and (b) and it is coded 1 for average sell recommendations (I/B/E/S score 1 or 2) in figures (c) and (d); it is otherwise coded 0. In figures (e) and (f), *Outcome* in week w is the difference between the median of individual recommendations for that week and the consensus forecast reported by I/B/E/S. $\Delta(Mean\ Uptalk)_{cftg}$ is the unexpected level of uptalk, as defined in Equation (5.2). X_{ft} is a vector of financial and stock characteristics and consensus analyst recommendations for firm f, and λ_f and θ_t are firm and time fixed effects, respectively. Standard errors are clustered by firm, and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

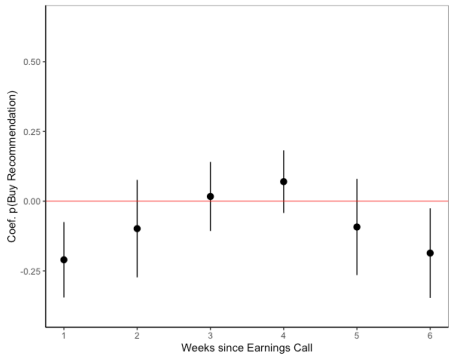
(a) Sell Recommendations: Female Executives



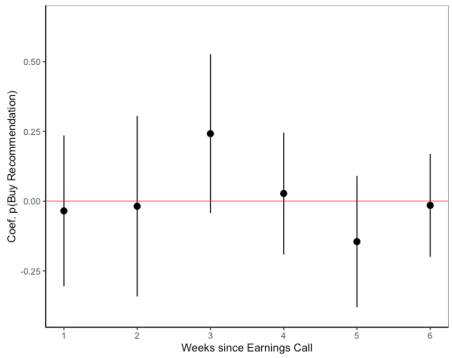
(b) Sell Recommendations: Male Executives



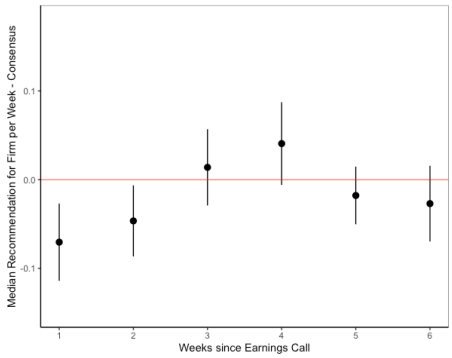
(c) Buy Recommendations: Female Executives



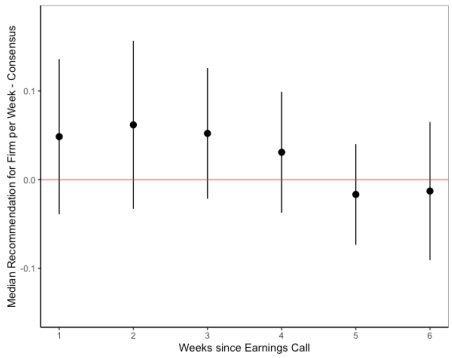
(d) Buy Recommendations: Male Executives



(e) Median - Consensus: Female Executives



(f) Median - Consensus: Male Executives



6

Analysts Downgrade Stock Recommendations When Female CEOs Use Uptalk

(Joint with Aharon Cohen Mohliver and Anantha Divakaruni[1])

Abstract

Despite evidence that female CEOs perform on par with men, women remain under-compensated and underrepresented in corporate leadership roles. One explanation for this gap is that they receive more negative evaluations. Yet, why and when negative evaluations occur remains underexplored. We find that analysts and investors respond negatively to newly appointed female CEOs who use high-rising intonation ('uptalk'), a female-typed speech pattern. In contrast newly appointed male CEOs who use 'uptalk' face no evaluative discount. This relationship between speech-type and evaluative discount was uncovered by applying a novel audio analysis method to a large dataset of the recordings of every earnings call surrounding CEO transitions in the US from 2011 to 2019. Our findings illuminate a novel driver of the evaluative discount faced by female leaders and demonstrate the unique value of audio analysis in understanding why some social groups are evaluated more negatively than their actual performance.

[1]The authors thank Hong Luo for his comments on an earlier draft, as well as audiences at the Strategic Management Conference 2023; the Strategy Science Conference 2023; the Academy of Management 2023; the HEC Paris Strategy Conference 2023; and at brown-bag seminars at London Business School.

6. Equity Analysts Downgrade Stock Recommendations When Female CEOs Use Uptalk 176

6.1. Introduction

Despite considerable efforts by policymakers, corporate insiders, investors, and social movements, women in leadership positions face, on average, less favourable evaluations and outcomes than men (Pletzer, Nikolova, Kedzior, and Voelpel, 2015; Carter, D'Souza, Simkins, and Simpson, 2010; Lee and James, 2007). In the US not only are women under-represented as CEOs, but also face lower stock recommendations from analysts, even when financial performance is comparable to that of male-led firms (Blau and Kahn, 2017; Pletzer, Nikolova, Kedzior, and Voelpel, 2015; Carter, D'Souza, Simkins, and Simpson, 2010; Lee and James, 2007).

As research on the disparities faced by female leaders grew, its focus shifted from mapping these gaps to understanding what drives them (Smith, Chown, and Gaughan, 2021; Begeny, Ryan, Moss-Racusin, and Ravetz, 2020; Berg and Lien, 2002). Until now, studies examining why analysts give lower stock recommendations for companies with Female CEOs have relied on transcripts of earnings calls, quarterly meetings where CEOs discuss their firms' financial results with analysts. Broadly, this work finds that female CEOs communicate in a more positive and less ambiguous tone and face more vague, negative, and aggressive questions from analysts (particularly male analysts) (De Amicis, Falconieri, and Tastan, 2021). Male analysts also tend to interrupt female CEOs more frequently (Brown, Crowley, and Elliott, 2020). Overall, these studies tend to find that communication style contributes more strongly to negative evaluations of female CEOs.

But transcripts are inherently limited in what measures can be derived from them. First, because the transcribed texts differ from the real content of communications (it does not include stutters, repetitions, gaps, or missing words), and second because they lack all the non-textual, verbal information such as the speaker's tone, pitch, speed, accent, diction, and a multitude of other highly informative vocal elements. Indeed, Divakaruni, Fritsch, Jones, and Morrison (2023) find that 'uptalk' (a high-rising tone at the end of declarative utterances) has a significant correlation with various financial indicators such as earnings, earnings forecasts, and bid-ask spreads, particularly for female CEOs.

We advance the idea that voice analysis can inform important questions about the evaluations faced by female leaders by using a unique dataset and novel analysis to explore whether an incom-

ing female CEO's uptalk influences analysts' downgrading of stock recommendations: unedited audio recordings of the earnings calls surrounding CEO transitions in U.S. publicly listed firms from 2011 to 2019. This rich dataset allows us to extract 15-second segments from presentations, questions, and answers by the 1,506 executives and 28,921 analysts. We measure 'uptalk' because of its strong association with female-typed speech and the continuous nature of its measure, which requires no assumptions about when speech becomes female-typed. By measuring how one speaks directly, rather than deducing it from transcribed content, we can more precisely capture gender-typed speech.

We ask whether the use of 'uptalk' by CEOs leads to changes in analyst recommendations, comparing the last three calls led by the outgoing CEO to the first four calls with the new CEO. This research design allows for implicit controls of non-observable variables related to both the firm and the CEO. In our regression models, for example, we use person-fixed effects and *company* × *person-fixed* effects that mitigate many potential omitted variable biases, including individual characteristics, managerial skill, and style, and selection in the choice of an incoming CEO.

Our study is among the first to leverage a large dataset of original voice recordings to examine how speech relates to the evaluation of female corporate leaders. This unique approach allows us to examine the relationship between gender-typed speech and analyst evaluations of female CEOs. We contribute to an emerging field of research that employs voice recordings instead of textual transcripts, joining only a handful of other studies that address gender issues through voice data (Seré, 2023; Talavera, Yin, and Zhang, 2022).

Our main finding is that equity analysts tend to downgrade recommendations for incoming female CEOs who use high levels of uptalk. This finding is robust even under stringent specifications that include firm, CEO, and call fixed effects. Analyst recommendations seem unaffected when men use low or high levels of uptalk, and when women use low levels of uptalk. Taken together, we conclude that the persistent downgrade in analyst recommendations for incoming female CEOs may be driven by a subtle interaction between gender and gender-typed speech, an explanation consistent with gender role norms.

We proceed to briefly describe our setting (financial analysts and earnings calls), then the data

and research design followed by a description of the results.

6.2. Literature Review

6.2.1. Gender-Typed Speech: Uptalk

Uptalk is defined as a rise in tonal pitch at the end of a declarative utterance compared to its body. There is a broad consensus in the sociolinguistics literature that uptalk is a female-typed speech characteristic (for review, see Brown, Hu, Shohfi, Francis, Zhang, and Xin, 2023; Lins, Roth, Servaes, and Tamayo, 2020; Chen, Nagar, and Schoenfeld, 2018). A higher incidence of uptalk has been observed among female speakers in a wide range of English-speaking communities (for review, see Cook, Esplin, Glass, Judd, and Olsen, 2019; Li, Sullivan, Xu, and Gao, 2013; Allen and Mendick, 2013; Henton, 1989; McConnell-Ginet, 1978). In fact, the association of uptalk with female speakers is so strong that research on transitioning transgender women finds that they increase their use of uptalk (Hancock, Colton, and Douglas, 2014; Hazenberg, 2013; Hazenberg, 2012). Given the strong association of uptalk and female-typed speech, we focus our main analysis on uptalk. For completion, we report corresponding text-based tonal measures that were used in prior literature (positive, negative, uncertain, financial, complex, and forward-looking tones in Appendix 6.A.1).

6.2.2. Equity analysts

Analysts play a pivotal role in financial markets, interpreting and mediating information for investors (Asquith, Mikhail, and Au, 2005). They conduct independent research, examine the firm's financial reports and disclosures, and participate in quarterly earnings calls. Analysts then issue a recommendation to investors on whether they should buy, sell, or hold the firm's stock. When investors act based on these recommendations, it directly impacts the stock price and the firm's value (Asquith, Mikhail, and Au, 2005). Previous research, relying on text data, has identified multiple elements shaping analysts' recommendations, including not only the content of earnings calls but also verbal and non-verbal cues like vagueness and temporal orientation (De Amicis, Falconieri,

and Tastan, 2021; DesJardine, Bansal, and Yang, 2019; Druz, Wagner, and Zeckhauser, 2016; Tetlock, Saar-Tsechansky, and Macskassy, 2008). Such non-verbal cues are particularly important when analysts encounter a CEO for the first time in their new role.

In existing literature, it is well-established that equity analysts are more likely to downgrade stock recommendations when a firm appoints a female CEO (Lee and James, 2007). Yet, there is no evidence that companies led by newly appointed female CEOs underperform. In fact, research suggests that female CEOs and CFOs often excel in areas that are particularly important to equity analysts, such as issuing more frequent and more accurate financial disclosures (Francoeur, Li, Singer, and Zhang, 2022; Barua, Davidson, Rama, and Thiruvadi, 2010; Peni and Vähämaa, 2010).

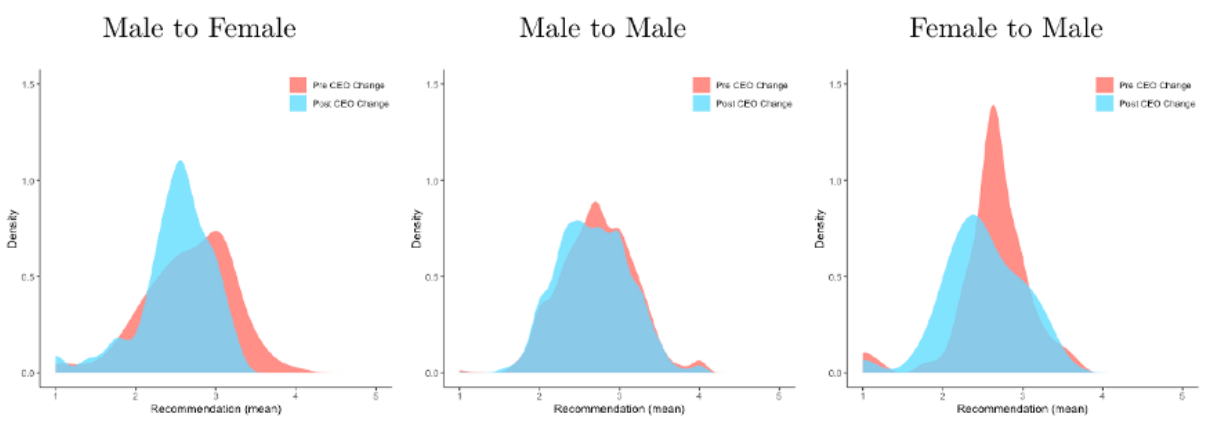
6.2.3. Analyst Reaction to Gender and Uptalk

Work in this field generally assumes that investors and analysts have the information required to assess a firm's future. Earnings calls reveal *surprise* - new information that could impact the company's stock price. Our main results indicate that the *surprise* is linked to uptalk by female CEOs. Figure 6.1's three graphs show the distribution of analyst recommendations after the first four earnings calls under a new CEO (blue), versus the last three calls with the outgoing CEO (red). The x-axis shows consensus recommendations, averaged on a 1-5 scale post-call, where 1 or 2 signifies *strong sell*, 3 is *hold*, and 4 or 5 is *strong buy*. The y-axis shows the density distribution of these recommendations.

Figure 6.1 shows a surprising symmetry: consensus recommendations by analysts become more positive when female CEOs are replaced by men, and more negative when male CEOs are succeeded by women. Importantly, these shifts in recommendations are not driven by outlier opinions. In fact, when the gender of the incoming CEO differs from that of the outgoing CEO, there is less variability in the consensus recommendations, pointing to a stronger alignment on a single consensus (buy/sell) recommendation.

To understand these shifts in recommendations during different-gender CEO transitions, we proceed to analyze gender-typed speech patterns in our recorded voice data. In other contexts, like call centers, speech patterns associated with gender and racial minorities have been shown to elicit

Figure 6.1. Distribution of consensus analyst recommendations around CEO transitions.
Figure 6.1 shows the distribution of consensus analyst recommendations following earnings calls among firms that underwent a CEO transition. Separate plots are shown for different types of gendered CEO transitions among firms. Within each plot, the red (blue) region depicts the distribution of consensus analyst recommendations before (after) the CEO change. Sample does not include firms that underwent interim changes in CEO, firms that did not have a CEO for one or more quarters during a transition, and firms that underwent a female-to-female change in CEO. Consensus analyst recommendations are coded on a negative point scale ranging from a "strong buy" denoted by 5 to a "strong sell" denoted by 1.



varied responses (Asquith, Mikhail, and Au, 2005). We expect similar effects during earnings calls.

6.2.4. Earnings calls

Earnings calls are conference calls held periodically between the managers of a publicly traded company and investment analysts, investors, and the media. They are an important source of information disclosure for investors, who often trade on the firm's stock based on this information, impacting its price, and through it the firm's value (Chen, Matsumoto, and Rajgopal, 2011; Hollander, Pronk, and Roelofsen, 2010).

During these calls, the firm's senior management, most often the CEO and CFO, initially present an overview of the company's performance and news they wish to share. This usually covers the firm's quarterly and annual financial reports, press releases, and other pertinent news. Afterward, they answer questions from participants (Blau, DeLisle, and Price, 2015). This unscripted Q&A dynamic can uncover new information. Sophisticated participants like analysts, can interpret nuances in answers (like evasion, diversion pauses, and tonal changes).

Transcripts of earning calls are commonly used in economics (Blau, DeLisle, and Price, 2015),

sociology (Vicinanza, Goldberg, and Srivastava, 2023) and social psychology (Li, Chiu, Kong, Cropanzano, and Ho, 2021) as they provide large data of detailed, uncensored discussions of firm strategy and performance. These transcripts can be easily matched to clearly identified and measurable dependent variables like post-call analyst recommendations and changes in the stock price, available in databases like I/B/E/S and CRSP. Prior research has examined shifts in these outcomes based on textual proxies of *tone* (see Figure 6.A.1 in the Appendix).

Text-based tonal measures derived from earnings calls have been used to examine how deception (Larcker and Zakolyukina, 2012), non-answers (Gow, Larcker, and Zakolyukina, 2021), political risk (Hassan, Hollander, van Lent, and Tahoun, 2019), commonality of CEO first name (Even-Tov, Huang, and Trueman, 2020), and conversation tone (Chen, Nagar, and Schoenfeld, 2018) impact the stock price. While this corpus of research uses transcripts of earnings calls, a notable exception is (Mayew and Venkatachalam, 2012) who uses a dataset of 615 audio recordings of some earnings calls that took place in a single year (2007) to look at how managerial affective states impact future firm performance. While this early study relies on a very small call sample and uses an ad-hoc threshold to distinguish positive from negative managerial tone, it suggests that audio recordings can be used to extract novel insights on managerial behaviour.

6.3. Data Description

Our initial sample comprises $60,038$ earnings calls covering $3,465$ US publicly listed firms from 2011 to 2019, featuring $203,159$ executives and $266,340$ analysts. The dataset contains over 2.29 million uptalk measurements. We manually identified 753 CEO transitions within these firms. Table 6.1 shows gender differences in these transitions: 671 with male-to-male CEOs, 50 with female-to-male CEOs, 30 male-to-female CEOs, and only 2 female-to-female CEOs. The latter is excluded due to its small number.

For each firm experiencing CEO transition, we use seven earnings calls: the last three under the outgoing CEO and the first four under the new CEO. Call transcripts (in text form) were sourced from Refinitiv StreetEvents and voice recordings from EarningsCast. For each earnings call, we then created an alignment map between the textual part of the speech of each participant (i.e. pre-

Table 6.1. Number of CEO change events
Table 6.1 presents the number of firms that underwent a CEO transition, grouped by the genders of the outgoing and incoming CEOs.

Differences in gender around CEO changes	Number of CEO changes
Male to Male	671
Male to Female	50
Female to Male	30
Female to Female	2

sentation, question, or answer) and its corresponding audio fragment, and split each participant's speech into 15-second segments. We measured the sound intensity (in decibels) of each audio segment using the *Praat* software. The uptalk of each participant's speech is calculated using these sound intensity measurements in the following way:

$$Uptalk = \begin{cases} \frac{Intensity_K}{K-1} - 1, & if\ Intensity_K > \frac{1}{K-1}\sum_{k=1}^{K-1} Intensity_k \\ 0, & otherwise \end{cases} \tag{6.1}$$

where K denotes the total number of 15-second audio segments in the speech of a given participant. This yields a continuous measure of uptalk bound at 0. Since the presentation part of the call is scripted, we focus on uptalk incidence among CEOs during the unscripted Q&A portion of the call and.

We obtained data on stock recommendations issued by analysts from I/B/E/S and, following prior work (Loh, 2011), coded it on a discrete five-point scale ranging from 1 (*strong sell*) to 5 (*strong buy*). A *buy* recommendation, therefore, takes the value of 1 in our regression models if the recommendation is 5 (*strong buy*) or 4 (*buy*). Similarly, a *sell* recommendation is a dummy equal to 1 if the recommendation is 1 (*strong sell*) or 2 (*sell*). A recommendation value of 3 indicates *hold*.

Lastly, data on the financial and stock characteristics of each firm were retrieved from Compustat and CRSP, respectively. Table 6.A.1 in the Appendix describes the variables used in our models.

6.4. Research Design

In any study evaluating CEOs, a significant challenge is the non-random matching of CEOs—especially female CEOs—to firms. To mitigate this issue to the best possible extent, our research design adopts an event study framework, in line with prior studies. We focus on earnings calls around CEO transitions, comparing analyst recommendations from the first four calls led by the incoming CEO to the last three calls led by the outgoing CEO.

Calls indexed from $t = -3$ to $t = -1$ took place under the outgoing CEO and calls indexed from $t = 0$ to $t = 3$ are under the new CEO. Our main regression specification is:

$$Y_{cpt} = \sum_{t=-3}^{t=3} \beta_t Call_{ct} + \sum_{t=-3}^{t=3} \gamma_t Uptalk_{cpt} + \sum_{t=-3}^{t=3} \pi_t Call_{ct} \times Uptalk_{cpt} + X_{ct} + \lambda_c + \mu_{cp} + \theta_t + \epsilon_{cpt} \quad (6.2)$$

where Y_{cpt} is the consensus mean of all recommendations issued by analysts on the firm's stock within the first 60-days following the call. $Call_{ct}$ is a set of dummy variables that take a value of one in the t-th call before or after the CEO transition, with $t = 0$ denoting the first call that took place under the new CEO. $Uptalk_{cpt}$ is the mean level of uptalk detected in the responses of CEO p of company c when answering analyst questions. X_{ct} denotes a vector of control variables capturing observable financial and stock characteristics of the firm at the end of each quarter. λ_c, μ_{cp} and θ_t denote firm (c), firm-CEO person (cp), and year-quarter (t) fixed-effects, respectively. The inclusion of fixed effects accounts for CEO traits and time-invariant firm attributes. This allows us to control for unobservable characteristics known to both firm and CEO, mitigating potential confounding factors in out analysis.

We run separate regressions for each of the three CEO transition categories: incoming and outgoing CEOs both male, female replacing male, and male replacing female. To reduce firm-level selection biases, we exclude firms with interim CEOs or those without a CEO for a quarter or more. We also do not conduct the analysis on firms in which both incoming and outgoing CEOs were female because this occurred only twice in our sample (see Table 6.1).

We use an event study design around the CEO transition date. By controlling for financial

performance and stock characteristics of the firm before and after the transition, we argue that there are no observable firm-specific factors that would cause firms to choose a male versus a female CEO. In other words, the glass cliff argument–whereby women are promoted to executive positions during times of crisis or during a recession when the chance of failure is higher–does not apply here. While it is reasonable to assume that the executives are unaware of the list of analysts who will be participating ahead of the call, we also control for the gender composition of analysts as this could affect the intensity of uptalk use. Unobservables are assumed to be random and are captured by the ϵ term. While we cannot control for firm-specific shocks as they are unobservable, our event study design averages across different firms undergoing CEO transitions at different time periods providing statistical validity to our estimates. We are careful not to attribute causality to our findings based on these estimates but rather provide aggregate comparisons within firms that undergo transitions in CEOs.

6.5. Results

6.5.1. Descriptive call features

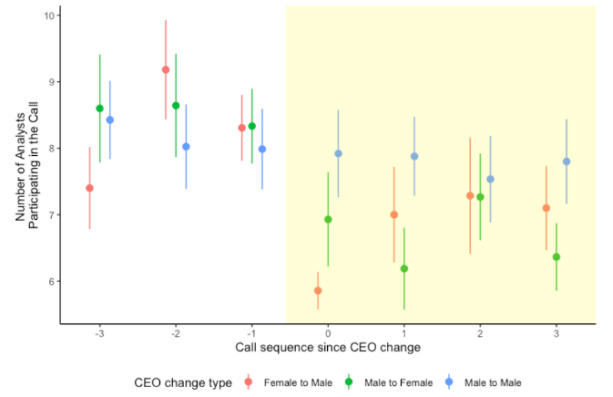
We first report descriptive statistics for each earnings call around a CEO transition event. In Figures 6.2, 6.3, 6.4, and 6.5, we follow the same convention: the x-axis in each plot represents the timing of the call (from call -3 to -1 before the CEO transition, shaded in white, to the first four calls under the new CEO, shaded in yellow). The color of the point estimates represents differences in the gender of the old and new CEOs (male-to-male CEO transitions in blue, male-to-female CEO transitions in green, and female-to-male CEO transitions in red) and the vertical lines represent ± 1 standard deviation from the mean.

In Figure 6.2, Panel (a) displays the number of analysts participating in each call, while Panel (b) shows the percentage of female analysts involved. For transitions between male CEOs (depicted in blue), both metrics exhibit little variation—the average number of analysts per call remains stable at approximately eight. When a male CEO succeeds a female CEO (indicated in red), the analyst count initially drops to six on the first call but rebounds to about 7.5 in subsequent calls. In contrast,

when female CEOs succeed male CEOs, the average number of analysts on the call decreases from eight prior to the CEO transition to fewer than seven during the first call led by the incoming female CEO. This decline continues, with the analyst count hovering between 6 and 7.5 in subsequent calls. Panel (b) in Figure 6.2 presents an even more pronounced trend: when a male CEO succeeds either a male or female predecessor, the proportion of female analysts participating in the call either drops or remains low. Conversely, when a female CEO takes over from a male CEO, the share of female analysts increases. Though this increase is most pronounced during the first call, the percentage of female analysts continues to be higher through the following three calls.

Figure 6.2. Analyst participation in earnings calls around CEO transitions.
Figure 6.2 presents changes in the extent of analyst coverage during earnings calls among firms that underwent a CEO transition. For each such firm, we consider the three most recent earnings calls under the outgoing CEO change and four subsequent calls under the incoming CEO (denoted by the yellow region). Call 0 on the x-axis denotes the first call that took place after the firm underwent a CEO change. Colours of the point estimates and confidence intervals indicate differences in the gender of the outgoing and incoming CEOs (male-to-male CEO transitions in blue, male-to-female CEO transitions in green and female-to-male CEO transitions in red). Sample does not include firms that underwent interim changes in CEO, firms that did not have a CEO for one or more quarters during a transition, and firms that underwent a female-to-female change in CEO. Vertical lines represent ±1 standard deviation in the outcome variable from its mean.

(a) Number of analysts participating in the call



(b) Percentage of analysts participating in the call that are female

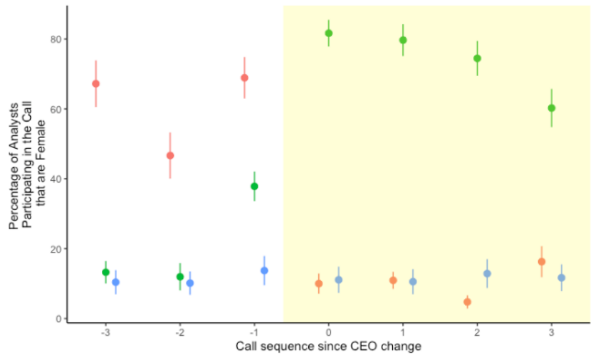
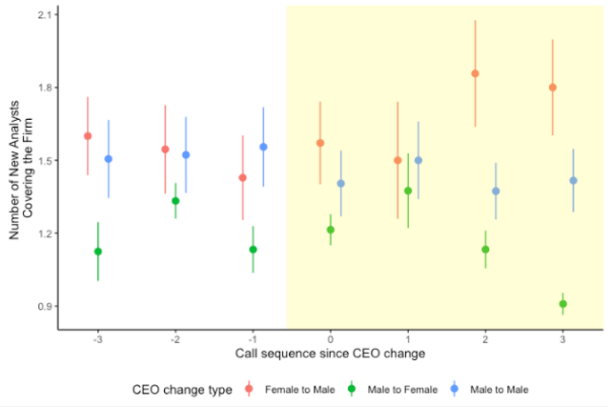


Figure 6.3(a) examines whether *new* analysts started to participate in the firm's earnings calls after a CEO transition. When men succeed women CEOs (red), more new analysts start participating in the company's calls. When women replace male CEOs (green), fewer newly participating analysts participate in the call. To get a better sense of these trends we plot in Figure 6.3(b) the number of new analysts scaled by the total number of analysts that issued a recommendation in the 180-day period preceding the call. The increase in coverage for men replacing female CEOs

persists even after scaling by the size of the analyst base, but the decline in coverage when women replace male CEOs does not.

Figure 6.3. Changes in analyst coverage around CEO changes within firms.
Figure 6.3 presents changes in the extent of new analyst coverage during earnings calls among firms that underwent a CEO transition. For each such firm, we consider the three most recent earnings calls under the outgoing CEO change and four subsequent calls under the incoming CEO (denoted by the yellow region). Call 0 on the x-axis denotes the first call that took place after the firm underwent a CEO change. Colours of the point estimates and confidence intervals indicate differences in the gender of the outgoing and incoming CEOs (male-to-male CEO transitions in blue, male-to-female CEO transitions in green and female-to-male CEO transitions in red). Sample does not include firms that underwent interim changes in CEO, firms that did not have a CEO for one or more quarters during a transition, and firms that underwent a female-to-female change in CEO. Vertical lines represent ±1 1 standard deviation in the outcome variable from its mean.

(a) Number of new analysts



(b) Percentage of analysts participating in the call that are female

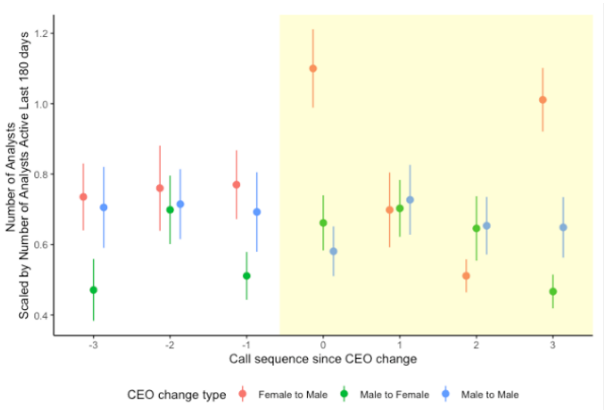
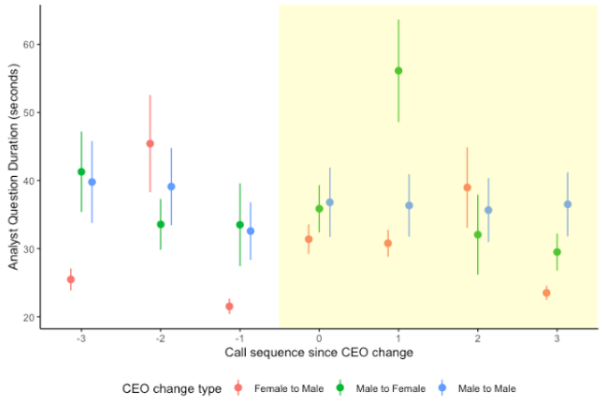


Figure 6.4 plots key statistics on the duration of questions and responses during the calls. Figure 6.4a looks at the duration of questions posed by analysts and Figure 6.4b the duration of responses by CEOs. As expected, both question and response duration increase during the first two calls under the incoming CEO, though the effect is larger and more persistent when men replace women as CEOs (6.4a, red).

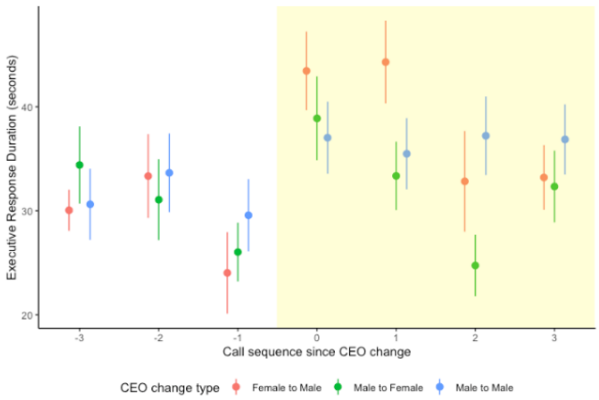
Since our analysis focuses on uptalk as a stereotypical gender-typed speech pattern, Figure 6.5 depicts the within-call (Panel (a)) and across-calls (Panel (b)) variation in uptalk from the last three calls of the outgoing CEO to the subsequent four calls of the firm under the succeeding CEO. We regress the mean level of uptalk of a CEO measured during each call on individual call sequence dummies ($t = -3$ to $t = 3$) and a vector of firm, firm-CEO, and year-quarter fixed, and plot the estimated coefficient of each dummy along with its 95% confidence interval. Separate regressions are run for each gender-CEO transition type.

Figure 6.4. Speech duration in earnings calls around CEO changes within firms.
Figure 6.4 shows changes in the duration of analyst and executive speeches during earnings calls among firms that underwent a CEO transition. For each such firm, we consider the three most recent earnings calls under the outgoing CEO change and four subsequent calls under the incoming CEO (denoted by the yellow region). Call 0 on the x-axis denotes the first call that took place after the firm underwent a CEO change. Colors of the point estimates and confidence intervals indicate differences in the gender of the outgoing and incoming CEOs (male-to-male CEO transitions in blue, male-to-female CEO transitions in green and female-to-male CEO transitions in red). Sample does not include firms that underwent interim changes in CEO, firms that did not have a CEO for one or more quarters during a transition, and firms that underwent a female-to-female change in CEO. Vertical lines represent ±1 standard deviation in the outcome variable from its mean.

(a) Analysts asking questions



(b) Executives responding to questions



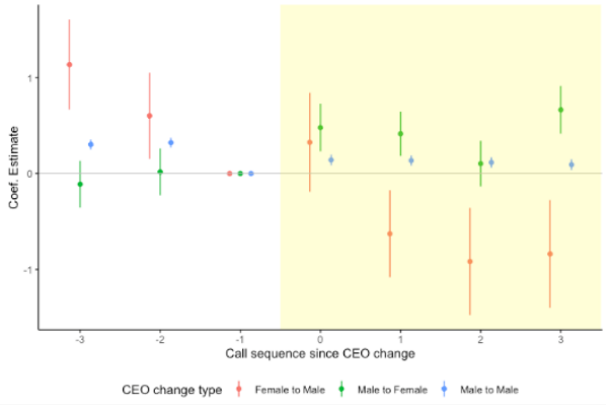
The patterns in these plots are striking. Both Figure 6.5a – within call – and Figure 6.5b – across calls – suggest significant changes in uptalk during calls following a CEO transition, but only when the genders of the outgoing and incoming CEOs differ. Specifically, there is a significant decrease in uptalk when men replace female CEOs. In contrast, uptalk increases when female CEOs replace men. Notably, there are no significant changes in the levels of uptalk when men replace men. There are no significant pre-trends in Figure 6.5b, depicting the across-call variation in mean uptalk among replaced CEOs. There is, however, evidence of a pre-trend in within-call variation in uptalk in the case of male-to-female CEO transitions. This is expected because uptalk varies significantly within a call depending on prevailing factors such as the nature of questioning (scrutiny by analysts and investors), CEO mental stress, and the CEO's willingness to reveal information when responding to questions.

Figure 6.5. Variation in uptalk around CEO changes within firms.
Each figure plots the estimated coefficient of uptalk incidence among CEOs during earnings calls among firms that underwent a transition in the CEO using the following event study specification:

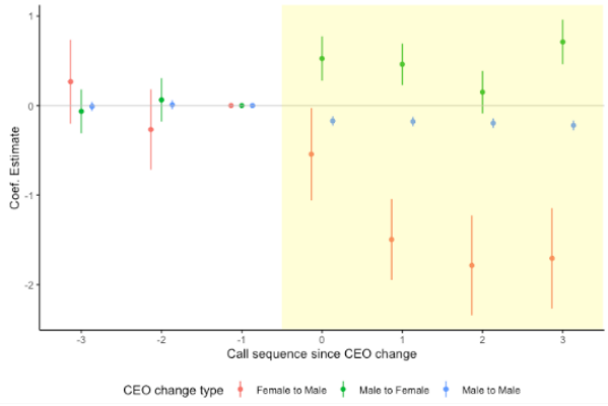
$$Y_{cpt} = \sum_{t=-3}^{t=3} \beta_t Call_{ct} + \sum_{t=-3}^{t=3} \gamma_t Uptalk_{cpt} + \sum_{t=-3}^{t=3} \pi_t Call_{ct} \times Uptalk_{cpt} + X_{ct} + \lambda_c + \mu_{cp} + \theta_t + \epsilon_{cpt}$$

where $Uptalk_{cpt}$ is the mean incidence of uptalk in the responses of CEO p of company c. $Call_{ct}$ is a set of dummy variables that take a value of one in the t_{th} call before or after the CEO change, where $t = 0$ denotes the first call that took place after the firm underwent a CEO transition. For each firm in our sample, we consider the three most recent earnings calls preceding the CEO transition and four subsequent calls under the incoming CEO (denoted by the yellow region). Sample does not include firms that underwent interim changes in CEO, firms that did not have a CEO for one or more quarters during a transition, and firms that underwent a female-to-female change in CEO. The regressions include firm c, year-quarter t, and CEO person p fixed effects. Standard errors are double-clustered by firm and person. Vertical lines represent 95% confidence intervals of the estimated coefficients.

(a) Within-call variation in uptalk among CEOs



(b) Across-call variation in uptalk among CEOs



6.5.2. Regression analysis

We proceed to run a linear regression model based on Equation 6.2, regressing the mean value (consensus) of all individual recommendations issued by analysts to the firm in the 60-day period following a call on dummy variables for each of the seven calls surrounding a CEO transition, CEO uptalk, and interaction terms between the call dummies and CEO uptalk. In Table 6.6, Columns 1-2 present regression estimates for male-to-male CEO transitions, Columns 3-4 for male-to-female CEO transitions, and Columns 5-6 for female-to-male CEO transitions. Every specification includes controls for the percentage of female executives and analysts participating in the call, the firm's financials characteristics (size, profitability, financial constraints, return on assets) and recent stock performance (return momentum, volatility, quarterly earnings per share (EPS), and the year-on-year change in quarterly EPS), and the consensus estimate of analyst recommendations

issued in the 90-days preceding the call. The scope and granularity of our sample allow us to run very tightly specified models incorporating fixed effects for individual CEOs and interaction terms between individual CEOs and their companies. This approach yields estimates that, while still correlational, approach causal validity, as many firm and individual attributes stay constant during these brief transitional periods. While the gender of the incoming CEO is known ahead of the call, analysts are unlikely to know the precise level of uptalk the CEO would use during the call, thus providing more strength to the validity of our estimates. Our results showing that a higher incidence of uptalk during a call result in a reduction in recommendations after the call supports the idea that uptalk is indeed unexpected, ex-ante.

Table 6.6 presents the corresponding results. Coefficients of $Call_{ct}$ in Models (1), (3), and (5) are generally not significant, both statistically and by economic magnitude, suggesting that analysts do not react strongly to replacement of the CEO itself or corresponding gender differences between the new and outgoing CEOs. The coefficient of $Uptalk_{cpt}$ is also not significant in any of these models, suggesting analysts do not react to the use of uptalk in of itself. However, interaction terms between $Call_{ct}$ and $Uptalk_{cpt}$ in Column (4) has negative and highly significant coefficients only for calls held under the incoming CEO (i.e. $t \geq 0$). This finding suggests a strong downward trend in consensus analyst recommendations when firms replace their male CEO with a female, but predominantly when the newly appointed female CEO uses high levels of uptalk in her responses to analyst questions. A one unit increase in uptalk incidence among female CEOs, lowers recommendations by 0.12 to 0.20 units during the first four earnings calls. A one standard deviation increase in uptalk among incoming female CEOs lowers the mean of consensus recommendation estimates towards a *sell* by between $0.107 (0.12 \times 0.89)$ and $0.18 (0.20 \times 0.89)$ units (see Table 6.A.3 in the internet appendix). However, the interaction terms between $Call_{ct}$ and $Uptalk_{cpt}$ for calls held under an outgoing male CEO in Column (4) (i.e. $t < 0$) show no such effect on recommendations if there is uptalk in the voice of these CEOs. Columns (2) and (6) show that uptalk has no significant relationship with analyst recommendations when the gender of the incoming and outgoing CEO is the same or involves a transition from a female to male CEO, respectively.

Figure 6.6. CEO changes, uptalk, and consensus in analyst recommendations

Table 6.6 presents regression estimates of the relationship between CEO transition and uptalk incidence among CEOs on consensus recommendations issued by analysts based on the following specification:

$$Y_{cpt} = \sum_{t=-3}^{t=3} \beta_t Call_{ct} + \sum_{t=-3}^{t=3} \gamma_t Uptalk_{cpt} + \sum_{t=-3}^{t=3} \pi_t Call_{ct} \times Uptalk_{cpt} + X_{ct} + \lambda_c + \mu_{cp} + \theta_t + \epsilon_{cpt}$$

where Y_{cpt} is the consensus mean of all recommendations issued by analysts on the firm's stock within the first 60-days following the call. $Call_{ct}$ is a set of dummy variables that take a value of one in the tth call before or after the CEO transition, with $t = 0$ denoting the first call that took place under the new CEO. For the estimates of $Call_{ct}$, $t = -1$ is set as the reference category. $Uptalk_{cpt}$ is the mean level of uptalk detected in the responses of CEO p of company c when answering analyst questions. X_{ct} denotes a vector of control variables capturing observable financial and stock characteristics of the firm at the end of each quarter. λ_c, μ_{cp}, and θ_t denote firm (c), firm-CEO person (cp), and year-quarter (t) fixed-effects, respectively. Standard errors double clustered by firm and CEO person are reported in parentheses. ***, **, and *denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Recommendation (mean)					
Call Type	Male to Male		Male to Female		Female to Male	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Call (t-3)	0.0291***	0.0237	0.0194	0.0119	0.0276	0.0541
	(0.0090)	(0.0151)	(0.0292)	(0.0634)	(0.0323)	(0.0553)
Call (t-2)	0.0147	0.0022	0.0334	0.0786**	-0.0296	-0.0039
	(0.0090)	(0.0151)	(0.0246)	(0.0304)	(0.0321)	(0.0419)
Call (t)	-0.0020	-0.0277	0.0448	0.0932	-0.0460	-0.0411
	(0.0221)	(0.0253)	(0.0610)	(0.0678)	(0.0931)	(0.1315)
Call (t+1)	0.0077	0.0085	0.0669	0.1088	-0.1481	-0.1339
	(0.0215)	(0.0236)	(0.0731)	(0.0686)	(0.0922)	(0.1112)
Call (t+2)	0.0174	-0.0043	0.0749	0.1039	-0.0996	-0.0523
	(0.0215)	(0.0212)	(0.0734)	(0.0694)	(0.0858)	(0.1027)
Call (t+3)	0.0016	-0.0102	0.0640	0.1428*	-0.0607	-0.0816
	(0.0213)	(0.0231)	(0.0727)	(0.0722)	(0.0800)	(0.1138)
Mean Uptalk (CEO)	-0.0038	-0.0252	-0.0301	0.0096	0.0267	0.0779
	(0.0067)	(0.0175)	(0.0204)	(0.0477)	(0.0233)	(0.0545)
Mean Uptalk (CEO) × Call (t-3)		0.0122		-0.0143		-0.0680
		(0.0257)		(0.1209)		(0.0819)
Mean Uptalk (CEO) × Call (t-2)		0.0270		-0.1398		-0.0683
		(0.0273)		(0.1949)		(0.0892)
Mean Uptalk (CEO) × Call (t)		0.0559**		-0.1627**		-0.0427
		(0.0260)		(0.0728)		(0.0833)
Mean Uptalk (CEO) × Call (t+1)		-0.0032		-0.1469***		-0.0635
		(0.0222)		(0.0552)		(0.0824)
Mean Uptalk (CEO) × Call (t+2)		0.0448*		-0.1204*		-0.1530**
		(0.0232)		(0.0634)		(0.0701)
Mean Uptalk (CEO) × Call (t+3)		0.0239		-0.2025***		0.0443
		(0.0226)		(0.0675)		(0.1029)
Female Executives (%)	-0.0607	-0.0597	0.1045	0.0761	-0.0414	-0.0515
	(0.0458)	(0.0456)	(0.0892)	(0.0845)	(0.1079)	(0.1115)
Female Analysts (%)	0.0123	0.0101	-0.0639	-0.0483	-0.0660	-0.0636
	(0.0251)	(0.0250)	(0.0521)	(0.0510)	(0.0746)	(0.0715)
Mean Uptalk (Analysts)	0.0190***	0.0188***	-0.0036	-0.0079	0.0000	0.0000
	(0.0070)	(0.0069)	(0.0158)	(0.0172)	(0.0000)	(0.0000)
Log(Assets)	0.0588*	0.0585*	0.0453	0.0458	-0.1355	-0.2089
	(0.0342)	(0.0346)	(0.0685)	(0.0636)	(0.2447)	(0.2610)
Profitability	0.1256	0.1113	0.6506	0.8881	0.9316	1.005
	(0.2217)	(0.2192)	(1.038)	(1.023)	(1.828)	(1.928)
Book/Equity	-0.0495*	-0.0496*	-0.1702*	-0.1361	-0.0193	-0.0427
	(0.0261)	(0.0261)	(0.0869)	(0.0995)	(0.0978)	(0.1047)
FinConstraints	-0.0007*	-0.0007*	-0.0010	-0.0015	-0.0033	-0.0041
	(0.0004)	(0.0004)	(0.0037)	(0.0040)	(0.0122)	(0.0127)
ROA	0.0260	0.0410	0.1598	0.1137	-3.613	-3.618
	(0.1785)	(0.1734)	(0.6973)	(0.7324)	(2.515)	(2.586)
SUE	0.0064**	0.0064**	-0.0008	-0.0005	0.0008	0.0022
	(0.0032)	(0.0032)	(0.0118)	(0.0119)	(0.0154)	(0.0161)
Momentum(-90,-2)	-0.0275	-0.0291	-0.0551	-0.0580	0.2682**	0.2568**
	(0.0180)	(0.0179)	(0.0556)	(0.0573)	(0.1160)	(0.1190)
Volatility(-90,-2)	0.4305	0.4650	-3.960**	-3.026*	-0.7182	-0.2947
	(0.4275)	(0.4269)	(1.717)	(1.715)	(2.159)	(2.288)
EPS	-0.0038*	-0.0040**	-0.0338*	-0.0387**	0.0103	0.0143
	(0.0020)	(0.0019)	(0.0171)	(0.0178)	(0.0403)	(0.0433)
$\Delta EPS^{\epsilon}_{t-4}$	0.0025**	0.0025**	0.0330*	0.0363*	0.0374*	0.0350*
	(0.0011)	(0.0011)	(0.0180)	(0.0191)	(0.0192)	(0.0204)
Last Recommendation (mean)	0.6390***	0.6395***	0.6392***	0.6253***	0.6586***	0.6595***
	(0.0378)	(0.0378)	(0.0851)	(0.0834)	(0.0348)	(0.0361)
Company x Person FE	✓	✓	✓	✓	✓	✓
Person FE	✓	✓	✓	✓	✓	✓
Observations	3,591	3,591	267	267	159	159
Adjusted R^2	0.8394	0.8393	0.8795	0.8770	0.8376	0.8094

6.5.3. Additional Analysis: Changes in Individual Analyst Recommendations

To assess the robustness of these findings, we repeat our analysis using a different set of outcomes capturing individual analyst recommendations. Our objective is to examine whether analysts change their own recommendations based on gender and uptalk. Due to variations in the recommendation of each analyst following an earnings call, it is not feasible to apply the primary specification defined in Equation 6.2 for this task, which is relevant only for consensus recommendation estimates that vary on a call-by-call basis. We, therefore, use the following specification:

$$Y_{cpta} = \sum_{t=-3}^{t=3} \beta_t Call_{ct} + \gamma Uptalk_{cpt} + X_{ct} + \lambda_c + \eta_a + \mu_{cp} + \theta_t + \epsilon_{cpta} \tag{6.3}$$

where Y_{cpta} denotes a recommendation issued by analyst a on the firm's stock within the first 60-days following the call. The remaining terms are exactly as specified in Equation (1), with the only additional term η_a denoting analyst fixed effects. We measure individual recommendations in two ways: (i) by differencing each recommendation from the consensus estimate over the same 90-day period after the call during which the recommendation was issued, and (ii) by differencing each recommendation from the one issued by the same analyst for the same firm in the preceding quarter. We run separate regressions for these two outcomes and report them in Panels (a) and (b) of Table 6.2.

The results of these analysis are remarkably consistent with the findings reported in Table 6.6 . Analysts reduce their own recommendations, only in the case of male-to-female CEO transitions, and particularly in relation to calls involving new female CEOs (as shown in Column (2) in both Panels). The incidence of uptalk during CEO responses serves to exacerbate this downtrend even further, as evidenced by the negative and statistically significant coefficient of mean uptalk in Column (2) of both Panels.

We complement the regression analysis with Figure 6.3[1], depicting density plots like the ones

[1]The distribution of analyst recommendations following the first four earnings calls held under a newly appointed female CEO (blue), compared to recommendations following the last three calls held under the outgoing male CEO (red). The x-axis represents the level of consensus recommendations, which is the average value of all recommendations issued by analysts on a 1-5 scale following an earnings call. A recommendation value of 1 or 2 denotes a "strong sell", 3 denotes "hold", whereas 4 and 5 indicate a "strong buy". The y-axis denotes the density distribution of recommendations.

presented in Figure 6.1, but where we split the sample to above and below mean levels of uptalk (Panels a and b respectively). Male-male CEO transitions are remarkably unaffected by the level of uptalk. But, high levels of uptalk by incoming female CEOs are clearly related to lower recommendations.

6.5.4. Additional Analysis: Investor Reaction to Uptalk

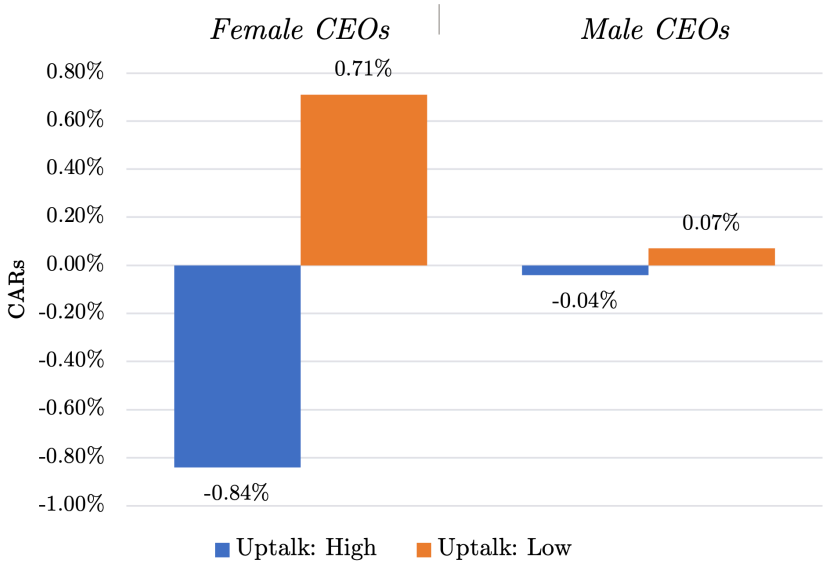
We repeat our analysis using another dependent variable: changes in the firm's stock price within one week of the call (Cumulative Abnormal Return, see Appendix 6.A.1 for a description of this measure). We choose CAR as a measure that is external to our dataset, where the CEO is evaluated by a different audience. Figure 6.7 presents the seven-day CAR for above and below mean levels of uptalk, in male and female CEO transitions.

The results are consistent with those relating to equity analysts' recommendations. Investors react strongly to the use of high uptalk by female CEOs. Interestingly, we find that investors do react to the use of uptalk by male CEOs, but the magnitude of their reaction is negligible: only -0.04% for high-levels of uptalk compared to -0.84% when female CEOs use high levels of uptalk. We observe a similar pattern for using low levels of uptalk: an increase of 0.07% for male CEOs and an increase of 0.71% for female CEOs. These effects are statistically significant both for men ($t = 3.57$) and women ($t = 13.00$), but they are economically meaningful only for female CEOs.

6.6. Discussion and Implications

Using voice data from 4,997 earnings calls around CEO transitions, we study how a CEO's gender-typed speech affects analyst stock recommendations. Our data's richness and size enable us to include CEO, firm, and call fixed effects in our models. We find that female CEOs using female-typed speech see stock recommendations decrease after their first earnings calls. Such speech type has no impact when men replace men, or when incoming female CEOs don't use female-typed speech. While not claiming causality, our design addresses many alternative explanations.

Figure 6.7. Abnormal Returns by Gender and Use of Uptalk.
Figure 6.7 shows the average cumulative abnormal returns (CARs) of firms that underwent a CEO transition during the first week following their earnings calls. The CARs are estimated over the period $t = -1$ to $t = 7$ where $t = 0$ denotes the day the earnings call took place. For each firm in the sample, we consider the three most recent earnings calls under the outgoing CEO change and four subsequent calls under the incoming CEO. The CARs are grouped by the gender of the incumbent CEO and by whether or not the incidence of uptalk during their participation in the earnings calls is above or below the sample average (into high and low uptalk, respectively). The sample does not include firms that underwent interim changes in CEO, firms that did not have a CEO for one or more quarters during a transition, and firms that underwent a female-to-female change in CEO.



Our findings clearly show that, controlling for firm, CEO, analyst, and call-level factors, equity analysts are more likely to downgrade stock recommendations for incoming female CEOs using high levels of uptalk. This validates our novel methodology leveraging earnings call voice recordings. Does this disparity in analyst reactions to uptalk by gender indicate an underlying bias, or could it suggest that uptalk by female CEOs is more informative? Either way, our results demonstrate the usefulness of analysing voice data, opening up new research avenues.

This chapter has some important implications for managerial practice and policy. We find that investors react strongly to the use of high uptalk by female CEOs and that the use of uptalk has important informational content. Investors want true information to be revealed as it allows them to better monitor their portfolio. From their perspective, having women rather than men steer their portfolio company's is optimal. While company's led by female CEOs don't underperform relative to company's led by male CEOs, their speech is more revealing of true performance and therefore reduces the magnitude of firm unobservables. This conclusion is compatible with the fact that

company's with female CEOs publish more voluntary disclosures (Nalikka, 2009). It is plausible that information is also revealed by how men speak but analysts are less sensitive to these variations, which is concerning given that most companies in our dataset our led by men.

6.A. Appendix

Figure 6.A.1. Textual tone in executive responses during earnings calls around CEO transitions within firms

Figure 6.A.1 presents changes in CEO sentiments among firms that underwent a CEO transition. We consider six different types of text-based sentiment measures: positive, negative, uncertain, financial, complex, and forward-looking. Each sentiment is measured using the Loughran-McDonald word lists that are publicly available. For example, positive sentiment in each earnings call is measured as the percentage of words spoken by the CEO during the call that are categorized as positive in the Loughran-McDonald word dictionary. For each such firm, we consider the three most recent earnings calls under the outgoing CEO change and four subsequent calls under the incoming CEO (denoted by the yellow region). Call 0 on the x-axis denotes the first call that took place after the firm underwent a CEO change. Colours of the point estimates and confidence intervals indicate differences in the gender of the outgoing and incoming CEOs (male-to-male CEO transitions in blue, male-to-female CEO transitions in green and female-to-male CEO transitions in red). Sample does not include firms that underwent interim changes in CEO, firms that did not have a CEO for one or more quarters during a transition, and firms that underwent a female-to-female change in CEO. Vertical lines represent ±1 standard deviation in the outcome variable from its mean.

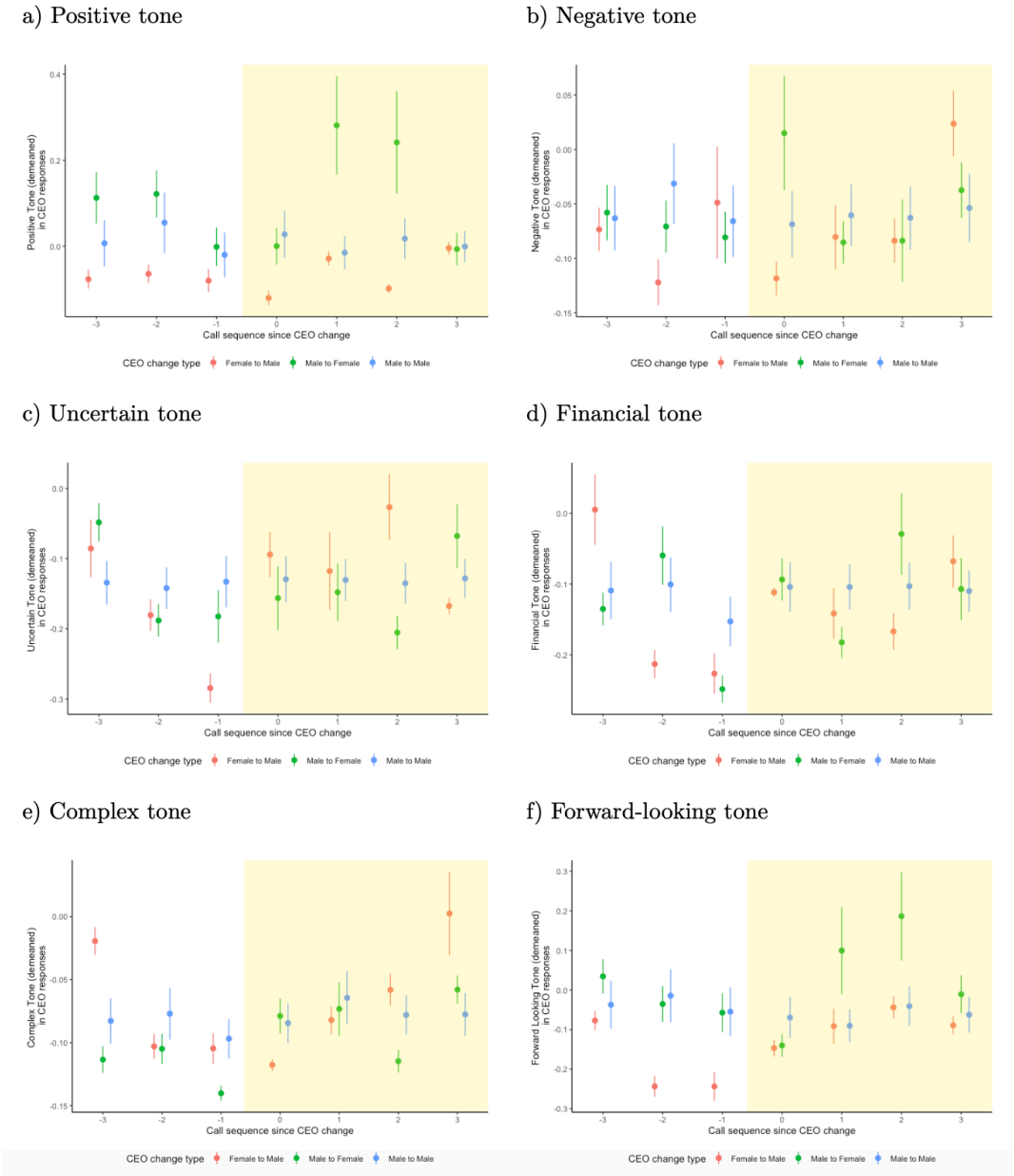


Figure 6.A.2. Distribution of CEO turnover by year and type

Figure 6.A.2 presents the number of firms undergoing CEO transitions per year during the sample period from 2011 to 2019. Each CEO transition is grouped depending on differences in the gender of the outgoing and incoming CEOs into male-to-male (yellow), male-to-female (green) and female-to-male (pink) types, respectively. Sample does not include firms that underwent interim changes in CEO, firms that did not have a CEO for one or more quarters during a transition, and firms that underwent a female-to-female change in CEO.

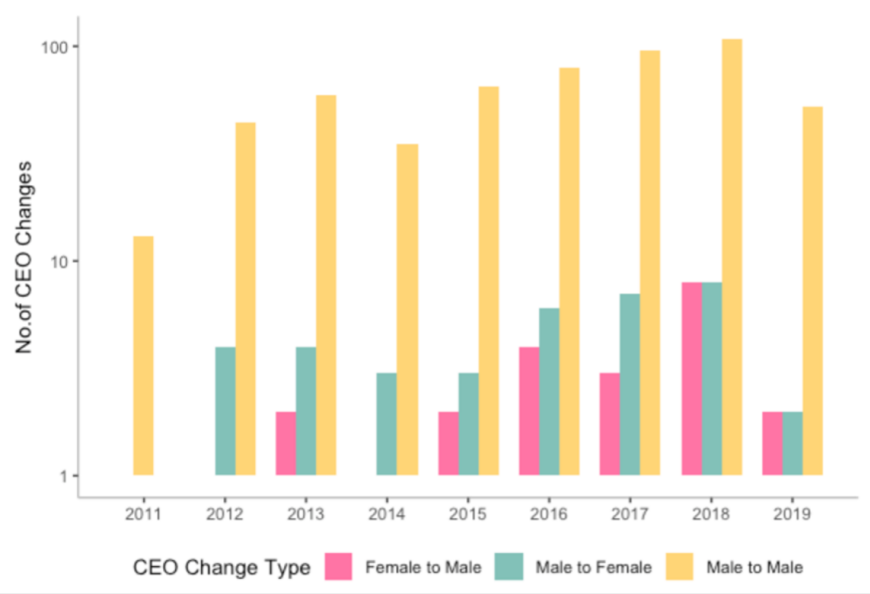


Table 6.A.1. Variable definitions

Variable	Definition
CEO Change	Seven dummy variables for each value of *t* indicating the first conference call of the incoming CEO, -3-0 are calls under the outgoing CEO, +1 - +3 are calls under the new (incoming) CEO.
Female Executives (%)	Percentage of executives participating in a call that are female.
Female Analysts (%)	Percentage of analysts participating in a call that are female.
Log(Assets)	Firm size measured in log of assets (in million USD).
Book/Equity	Firm book-to-equity ratio, calculated by dividing.
FinConstraints	Kaplan-Zingales financial constraints measure (Kaplan and Zingales, 1997).
ROA	Return on assets of the firm
SUE	Standardized unexpected earnings (Bernard and Thomas, 1990).
Momentum(-90,-2)	Return on the firm's stock over the 90-day period prior to the call.
Volatility(-90,-2)	Volatility on the firm's stock over the 90-day period prior to the call.
EPS	Firm's most-recent earnings per share.
$\Delta EPStt-4$	Change in the firm's earnings per share between the most recent quarter and four quarters prior.
Last Recommendation (mean)	Mean of the recommendations issued by analysts over the prior 90 day period to the call.

Table 6.A.2. Descriptive statistics: all earnings calls

Table 6.A.2 presents summary statistics for all earnings calls in our sample for firms that underwent a CEO transition. Sample does not include firms that underwent interim changes in CEO, firms that did not have a CEO for one or more quarters during a transition, and firms that underwent a female-to-female change in CEO. See Section 4 for a description of the sample.

	Mean	SD	Minimum	25 pct	Median	75 pct	Maximum
Call-level characteristics							
CEO Uptalk	0.59	0.75	0	0.20	0.40	0.71	10.47
CEO Response Duration (seconds)	41.74	47.72	0.04	4.48	29.24	62.20	1888.96
No. of Executives	2.78	1.12	1	2	3	3	11
Female Executives (%)	10.70	19.61	0	0	0	20	100
No. of Analysts	6.48	4	1	3	6	9	24
Female Analysts (%)	14.98	30.02	0	0	0	10	100
Analyst recommendations							
Recommendation (consensus)	2.67	0.47	1	2.35	2.67	3	4
Recommendation (net of consensus)	-0.01	0.10	-0.30	-0.09	0	0.06	0.40
Recommendation (net of previous by same analyst)	0	0.04	-0.30	0	0	0	0.30
Firm characteristics							
Assets ($ billions)	15.07	32.01	0.01	1.24	4.11	12.57	190.17
Book/Equity	0.54	0.48	-0.89	0.26	0.45	0.74	3.09
Profitability (%)	2.48	4.01	-30.30	1.14	2.73	4.13	12.74
ROA (%)	0.63	4.13	-38.71	0.23	1.09	2.09	10.69
Financial Constraints	2.84	7.71	-32.81	0.77	1.92	3.62	50.58
SUE	0.02	1.03	-2.47	-0.63	0.06	0.68	2.47
Momentum(-90,-2)	-0.01	0.21	-1.55	-0.11	-0.01	0.09	3.64
Volatility(-90,-2)	0.02	0.01	0	0.01	0.02	0.02	0.26
EPS	0.55	1.39	-25.37	0.09	0.45	0.91	55.20
ΔEPS^{q}_{t-4}	-0.01	2.20	-119.72	-0.13	0.04	0.22	58.72
No. of earnings calls				4,997			

Table 6.A.3. Descriptive statistics: earnings calls under male and female CEOs

Table 6.A.3 presents summary statistics for earnings calls in our sample under male and female CEOs. Our sample includes only firms that underwent a CEO transition, and does not include firms that underwent interim changes in CEO, firms that did not have a CEO for one or more quarters during a transition, and firms that underwent a female-to-female change in CEO. See Section 4 for a description of the sample. The last two Columns present the t-test statistics of the difference in means between characteristics observed during earnings calls under male and female CEOs, respectively.

	Earnings calls under						*t*-test
	Male CEOs			Female CEOs			Diff. in means
	Mean	SD	Median	Mean	SD	Median	*t*-stat (pval)
Call-level characteristics							
CEO Uptalk	0.59	0.74	0.40	0.69	0.89	0.45	8.64***
CEO Response Duration (seconds)	41.90	48.00	29.40	38.94	42.54	26.94	-5.12***
No. of Executives	2.79	1.12	3.00	2.52	0.81	2.00	-23.7***
Female Executives (%)	7.52	15.82	0	53.45	20.70	50	163.96***
No. of Analysts	7.74	4.15	7.00	7.13	4.34	6.00	-10.23***
Female Analysts (%)	9.18	22.38	0	86.88	21.29	100	265.64***
Analyst recommendations							
Recommendation (consensus)	2.70	0.44	2.70	2.55	0.44	2.58	-23.38***
Recommendation (net of consensus)	-0.01	0.10	0	0	0.09	0.03	6.34***
Recommendation (net of previous by same analyst)	0	0.04	0	-0.01	0.03	0	-7.63***
Firm characteristics							
Assets ($ billions)	14.82	32.12	4.08	19.20	29.92	5.24	10.54***
Book/Equity	0.54	0.48	0.45	0.50	0.45	0.44	-6.16***
Profitability (%)	2.45	3.97	2.75	2.86	4.61	2.45	6.45***
ROA (%)	0.61	4.12	1.10	0.94	4.26	1.06	5.65***
Financial Constraints	2.90	7.65	1.95	1.91	8.49	1.65	-8.42***
SUE	0.02	1.03	0.06	0.02	0.97	-0.02	0.09 (0.932)
Momentum(-90,-2)	0	0.21	-0.01	-0.02	0.18	-0.02	-5.57***
Volatility(-90,-2)	0.02	0.01	0.02	0.02	0.01	0.02	-16.07***
EPS	0.54	1.38	0.45	0.71	1.45	0.48	8.28***
ΔEPS^{q}_{t-4}	-0.01	2.25	0.04	0.03	1.27	0.04	2.26**
No. of earnings calls		4,740			257		

Table 6.2. CEO changes, uptalk, and changes in individual analyst recommendations

Table 6.2 presents regression estimates of the relationship between CEO transition and uptalk incidence among CEOs on individual analyst recommendations based on the following specification:

$$Y_{cpta} = \sum_{t=-3}^{t=3} \beta_t Call_{ct} + \gamma Uptalk_{cpt} + X_{ct} + \lambda_c + \eta_a + \mu_{cp} + \theta_t + \epsilon_{cpta}$$

where Y_{cpta} denotes the individual recommendations issue by analyst a on the firm's stock within the first 60-days following the call. $Call_ct$ is a set of dummy variables that take a value of one in the tth call before or after the CEO transition, with $t = 0$ denoting the first call that took place under the new CEO. $Uptalk_{cpt}$ is the mean level of uptalk detected in the responses of CEO p of company c when answering analyst questions. X_{ct} denotes a vector of control variables capturing observable financial and stock characteristics of the firm at the end of each quarter. λ_c, μ_{cp}, and θ_t denote firm (c), firm-CEO person (cp), and year-quarter (t) fixed-effects, respectively. Individual recommendations are measured in two ways: In Panel (a) by differencing each recommendation from the consensus mean estimate over the same 90-day period following the earnings call during which the recommendation was issued, and in Panel (b) by differencing each recommendation from the one issued by the same analyst for the same firm in the preceding quarter. Standard errors double clustered by firm and CEO person are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

(a) Differences from consensus estimate

Dependent Variable:	Recommendation (net of consensus estimate)		
Call Type	Male to Male	Male to Female	Female to Male
Model:	(1)	(2)	(3)
Call (t-3)	-0.0154	-0.1668	-0.0034
	(0.0228)	(0.1345)	(0.0451)
Call (t-2)	-0.0090	-0.0904	0.1754**
	(0.0234)	(0.1477)	(0.0709)
Call (t)	0.0078	-0.6483***	0.3683*
	(0.0379)	(0.1931)	(0.1920)
Call (t+1)	0.0009	-0.5829***	0.5111***
	(0.0388)	(0.1951)	(0.1692)
Call (t+2)	-0.0032	-0.7234***	0.4364***
	(0.0384)	(0.1784)	(0.1318)
Call (t+3)	0.0058	-0.4390*	0.2874*
	(0.0381)	(0.2262)	(0.1665)
Mean Uptalk (CEO)	-0.0048	-0.1546***	-0.1802***
	(0.0105)	(0.0512)	(0.0211)
Firm Financial/Stock Controls	✓	✓	✓
Most Recent Consensus Recommendation	✓	✓	✓
Company x Person FE	✓	✓	✓
Analyst FE	✓	✓	✓
Observations	63,129	3,241	2,635
Adjusted R^2	0.5983	0.9332	0.9527

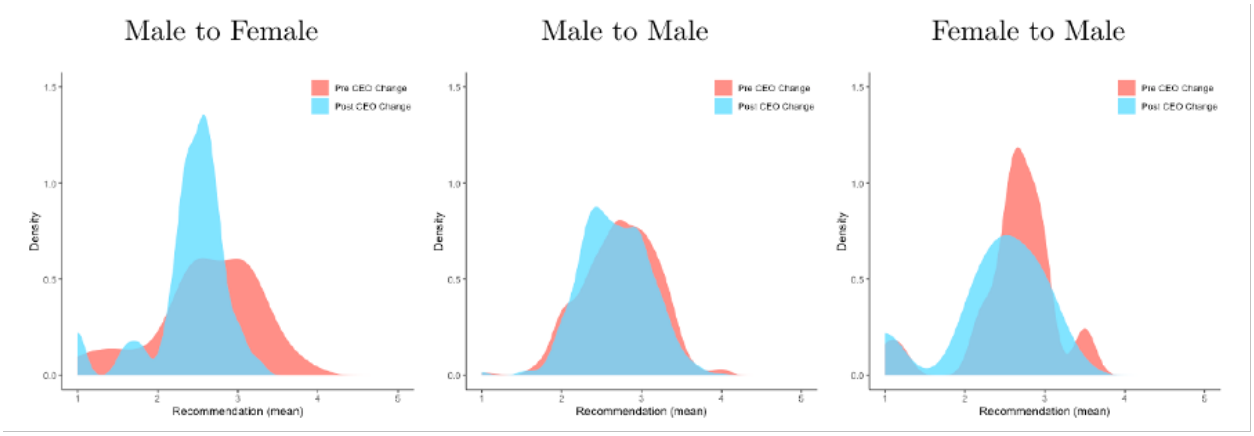
(b) Recommendation revisions from most recent estimate

Dependent Variable:	Recommendation revision (net of previous by same analyst)		
Call Type	Male to Male	Male to Female	Female to Male
Model:	(1)	(2)	(3)
Call (t-3)	-0.0087	-0.0789	0.0016
	(0.0100)	(0.1151)	(0.0401)
Call (t-2)	-0.0077	-0.0832	-0.0255
	(0.0112)	(0.1396)	(0.0427)
Call (t)	-0.0052	-0.0464	0.0813
	(0.0151)	(0.1328)	(0.0812)
Call (t+1)	-0.0054	0.0571	-0.0096
	(0.0152)	(0.1407)	(0.0886)
Call (t+2)	-0.0042	-0.2038*	0.0563
	(0.0146)	(0.1106)	(0.0646)
Call (t+3)	-0.0056	-0.1362	0.0838
	(0.0140)	(0.1698)	(0.0837)
Mean Uptalk (CEO)	0.0077	-0.1001***	-0.0333*
	(0.0063)	(0.0314)	(0.0172)
Firm Financial/Stock Controls	✓	✓	✓
Most Recent Consensus Recommendation	✓	✓	✓
Company x Person FE	✓	✓	✓
Analyst FE	✓	✓	✓
Observations	63,129	3,241	2,635
Adjusted R^2	0.3749	0.8311	0.7210

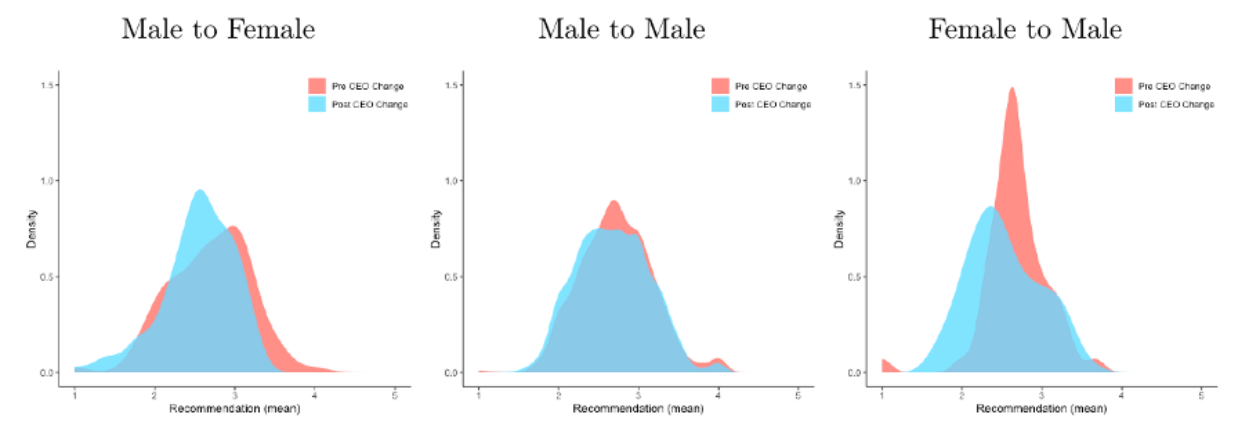
Table 6.3. Distribution of consensus analyst recommendations around CEO transitions

Figure 6.3 shows the distribution of consensus analyst recommendations following earnings calls among firms that underwent a CEO transition, conditional on the usage of uptalk by the CEO. Panel (a) shows consensus recommendations for earnings calls characterized by above-average levels of mean CEO uptalk, and Panel (b) for earnings calls characterized by below-average levels of mean CEO uptalk. Within each Panel, separate plots are shown for different types of gendered CEO transitions among firms. Within each plot, the red (blue) region depicts the distribution of consensus analyst recommendations before (after) the CEO change. Sample does not include firms that underwent interim changes in CEO, firms that did not have a CEO for one or more quarters during a transition, and firms that underwent a female-to-female change in CEO. Consensus analyst recommendations are coded on a negative point scale ranging from a "strong buy" denoted by 5 to a "strong sell" denoted by 1.

(a) Earnings calls with *high* uptalk (i.e. mean uptalk $\geq$ sample average)



(b) Earnings calls with *low* uptalk (i.e. mean uptalk $<$ sample average)



7

Conclusion

We conclude the thesis by briefly discussing some integrative findings and suggesting promising future research directions.

7.1. Integrative Reflections

In this section, we split the thesis into two thematic groups: language (Chapter 3 and Chapter 4) and gender (Chapter 5 and Chapter 6). We discuss them in turn.

7.1.1. Language

Language provides scaffolding for organizational life. The thesis argues that vocabularies generate community-specific understandings of the outside world. The meaning of a language exists *between* its speakers and not in mental images that the language attempts to convey. Both in Chapter 3 and Chapter 4, "meaning is use" (Wittgenstein, 1953). In the former, a process model links environmental cues to organizational actions where language mediates organizational adaptation and learning. Language underpins organizational routines, and these routines can cause the organization to become decoupled from means-end rationality. We argue that organizational responses to social and technological shocks are essentially linguistic and recategorization can cause discontinuous and unpredictable changes to organizational behaviour. In the latter, vocabularies are formed through the repeated performance of a task that requires an actor with an object to convey to others in the same organization in which category that object sits.

In both chapters, language, categorization, and organizational action interact to create persistent and path-dependent organizational behaviours. Linguistic rules become part of the taken for granted reality of the organization and so categorical understandings of external environment become wired into organizational procedures. In Chapter 3, categorical and conative meanings are created in distinct language games that are played in distinct communities so the process of categorization can evolve independently of the practical actions taken in response to that categorization. This causes organizational behaviour to remain relatively stable over time as organizations assign new environmental cues to pre-existing categories rather than creating new ones. This would only lead the categorical meaning to change without affecting its conative goal. Thus, organizations can process new cues by pursuing outdated goals. Similarly, Chapter 4 examines the ways in which vocabularies become imprinted on organizations. Through repeated task performance, experimental organizations converge on codes to identify specific images and those codes become fairly stable. It shows that, once imprinted, an organizational vocabulary has long-lived effects on the ways in which organizational performance. They also affect the way that organizational members perceive their world and, therefore, upon their actions. Moreover, we also find that founders imprint vocabularies on organizations that significantly outlast their membership. These vocabularies even affect the recall of members that joined the organization after all their founder had left. Taken together, the two chapters suggest that in order to change organizational responses to external cues, we have to start by understanding the way that organizational language works and intervening directly in its use. Specifically, we need to intervene in the way that language drives categorization both by directing attention and also by suggesting family resemblances.

7.1.2. Gender

The thesis argues that information is revealed not only by what people say but also by *how* they say it. Both Chapter 5 and Chapter 6 build on a methodological innovation which uses audio recordings to identify gendered speech patterns to study the way in which organizations are perceived by outside observers. Specifically, we introduce the concept of uptalk to the management and economics literature and develop a quantitative method to measure its use. We show that there are

measurable differences in how men and women speak and show that these patterns reveal information to markets. Markets respond differently to female uptalk, but this is not driven by unconscious biases. Rather, it is rational response because the use of uptalk contains important informational content. In Chapter 5, we demonstrate that the female-typed uptalk speech pattern influences analyst perceptions during earnings calls and, therefore affect their buy and sell recommendations. We find that analysts respond negatively when female executives use unexpectedly high levels of uptalk in earnings calls. Realised earnings are lower when women use more uptalk than usual so this response appears to be rational. Indeed, our data indicates that analysts under react to uptalk by female executives. We also find that market perceptions of risk increase when female executives speak in earnings calls. Post #MeToo, analyst reactions remained unchanged despite the prevalence of uptalk by female executives decreasing. Surprisingly, however, analysts began to view uptalk by male executives more favourably. That is, one of the unintended consequences of a movement aimed at achieving greater social justice for women was to reward men who exhibited female-typed speech patterns.

Similarly, in Chapter 6, we find that female CEOs using female-typed speech see the recommendations of their firms' stock decrease after the first set of earnings calls held under their tenure. Uptalk has no effect on transitions in which men replace male CEOs, nor in instances in which incoming female CEOs do not use female-typed speech. Taken together, these results contribute to research on social norms, showing that while men don't face negative feedback for adopting speech female-typed speech patterns, women are evaluated more critically for adhering to these gender-congruent speech norms and their use has informational content. This pattern suggests that, in the context quarterly earnings calls, gender role norms influence analysts' expectations over generalized social norms pertaining to speech style.

When women are in leadership positions, markets have better information about businesses and, hence, market-based incentives can more effectively steer organizational behaviours. Investors want true information to be revealed during quarterly earnings calls that allows them to better monitor their portfolio. This thesis shows that company's led by female executives don't underperform relative to companies led by male executives but the former's speech revealing of true

financial performance. This may or may not be a desirable outcome but it is an empirical fact. If preferences change when women are put in leadership positions, then market powers to incentivize also ought to change.

7.2. Future Research

In this section, we outline some promising future research directions that are opened up by the findings in this thesis.

7.2.1. Language

Chapter 3 and Chapter 4 open a number of questions on the use of language in organizations. First, we plan on empirically testing the hypotheses that we derive from the process model in Chapter 3. This would ideally be accomplished through comparative case studies to understand why organizations in the same market react differently to certain technological or socio-cultural shocks. We would need to look at their internal categorization process and associated goals. This analysis would be done using quantitative or qualitative methods.

Moreover, we plan on altering the experimental protocol used in Chapter 4 to study the impact of organizational structure on the formation, transmission, and persistence of vocabularies in organizations. The current set up is hierarchical and turnover follows a predictable pattern. Every new participant that enters the experiment joins as trainee of Organization 1 and all promotions follow the same pattern. Participants can anticipate that the team that performed the image recognition task in Organization 1 will also compose Organization 2 after three rounds. We would first like to test how a flat structure would compare with a hierarchical structure. Instead of constraining what the trainee and associate can do, we would allow both to speak and click the images. We would also like to test the impact of randomizing the staff turnover. In the current set up the director leaves the organization, the associate is promoted to director, and the trainee is promoted to associate at the end of every round. This predictable pattern could be anticipated by the participants. Alternatively, the role switches could be randomized so that participants could not anticipate the group

composition of future rounds. Furthermore, in the current experimental set up, the founders create codes that allow for efficient image identification. We would like to intervene in how vocabularies actually work by seeding the founding vocabulary. Holding the vocabulary constant, we would then investigate what factors explain superior organizational performance in later tasks.

In Chapter 4, the protocol is only used to investigate efficiency. We plan on adapting the protocol to a negotiation setting. Instead of having participants perform an image identification task, we could use widely established case studies in the negotiations literature (e.g. Sugar Bowl, Energetix vs. Generex, Yerba Mate, Solaris) where participants need to negotiate over a financial transaction instead. This opens a number of interesting theoretical questions such as: can negotiators' linguistic alignment at the beginning of an interaction predict negotiation outcomes? Does the speed with which negotiators linguistically align predict negotiation outcomes? How do developments in negotiators' linguistic alignment over the course of a deal predict the outcome? Furthermore, financial rewards in the current experimental set up are attached exclusively to group performance. In a future iteration of the study, we could induce a cooperative versus competitive orientation by instructing them to focus on individual versus collective outcomes. We could also manipulate the initial economic offers to be aligned (close to each other) or misaligned (far from each other).

Finally, the experimental protocol in Chapter 4 could also be altered to investigate ethical questions. The image identification task could be replaced with a matrix task commonly used in the behavioural ethics literature. We could then design an intervention where some organizations are founded using ethical language whereas other organizations are founded using unethical language. This would allow us to study whether the ethical stance of the founding team has persistent effects in the behaviour of participants in future rounds.

7.2.2. Earnings calls data

The methods developed in Chapter 5 and Chapter 6 could be used to investigate the real economic effects of cultured mores in other settings. A population can be divided culturally in many ways. When a particular cultural grouping is associated with a way of speaking, a database of potentially market-sensitive speech can be with data about cultural markers and prices to study the real effects

of that pattern. This type of analysis has the potential to shed new light on the real economic effects of perceptions not only of gender, but also of race, class, and educational achievements.

The earnings calls dataset could also be used investigate deceptive communications. The last few years has seen an increasing number of companies setting diversity and climate targets. Notwithstanding these targets, companies are being called out for diversity-washing and green-washing. Regarding diversity-washing, it would be interesting to document the discrepancies between companies' external stances on diversity, equity, and inclusion (DEI) and their hiring practices and then try to capture differences in communication styles between low-discrepancy and high-discrepancy companies during earnings calls. Regarding green-washing, this year has seen a number of lawsuits brought against companies that made misleading environmental claims. For example, Keurig Green Mountain reached a class-action settlement and agreed to pay $10 million after customers sued over its coffee pod recycling claims and Shell directors are being personally sued over their climate strategy, which claimainst say are inadequate to meet public climate targets. In response to these criticisms, the FCA and other regulatory authorities have also proposed new rules to tackle greenwashing. Similarly, it would be impactful to document the discrepancies between companies' net-zero targets and their actual climate-mitigation activities and then try to capture differences in communication styles between low-discrepancy and high-discrepancy companies during earnings calls.

We leave these questions for future research.

Bibliography

Adams, R. B. 2017. Boards, and the Directors Who Sit on Them. In *The Handbook of the Economics of Corporate Governance*, pp. 291–382. Elsevier.

Adams, R. B., and D. Ferreira. 2009. Women in the Boardroom and Their Impact on Governance and Performance. *Journal of Financial Economics* 94:291–309.

Adams, R. B., and P. Funk. 2012. Beyond the Glass Ceiling: Does Gender Matter? *Management science* 58:219–235.

Adams, R. B., and T. Kirchmaier. 2016. Women on Boards in Finance and STEM Industries. *American Economic Review* 106:277–81.

Adhikari, B. K., A. Agrawal, and J. Malm. 2019. Do Women Managers Keep Firms out of Trouble? Evidence from Corporate Litigation and Policies. *Journal of Accounting and Economics* 67:202–225.

Ahern, K. R., and D. Sosyura. 2015. Rumor Has It: Sensationalism in Financial Media. *Review of Financial Studies* 28:2050–2093.

Albanesi, S., and C. Olivetti. 2009. Home Production, Market Production and the Gender Wage Gap: Incentives and Expectations. *Review of Economic Dynamics* 12:80–107.

Alesina, A., B. Brioschi, and E. La Ferrara. 2016. Violence Against Women: A Cross-Cultural Analysis for Africa. *NBER Working Paper Series* .

Alesina, A., P. Giuliano, and N. Nunn. 2013. On the Origins of Gender Roles: Women and the Plough. *Quarterly Journal of Economics* 128:469–530.

Allen, K., and H. Mendick. 2013. Keeping It Real? Social Class, Young People and 'Authenticity'in Reality TV. *Sociology* 47:460–476.

Amburgey, T. L., and A. S. Miner. 1992. Strategic Momentum: The Effects of Repetitive, Positional, and Contextual Momentum on Merger Activity. *Strategic Management Journal* 13:335–348.

Anthony, S. 2016. Kodak's Downfall Wasn't about Technology. *Harvard Business Review* 15:1–5.

Argote, L. 2012. *Organizational Learning: Creating, Retaining and Transferring Knowledge*. New York, NY: Springer Science & Business Media.

Argote, L., and H. R. Greve. 2007. A Behavioral Theory of the Firm: 40 Years and Counting: Introduction and Impact. *Organization Science* 18:337–349.

Arrow, K. J. 1974. *The Limits Of Organization*. New York: Norton & Company.

Asquith, P., M. B. Mikhail, and A. S. Au. 2005. Information Content of Equity Analyst Reports. *Journal of Financial Economics* 75:245–282.

Atwater, L. E., A. M. Tringale, R. E. Sturm, S. N. Taylor, and P. W. Braddy. 2019. Looking Ahead. *Organizational Dynamics* 48:100677.

Audia, P. G., and H. R. Greve. 2006. Less Likely to Fail: Low Performance, Firm Size, and Factory Expansion in the Shipbuilding Industry. *Management Science* 52:83–94.

Audia, P. G., E. A. Locke, and K. G. Smith. 2000. The Paradox of Success: An Archival and a Laboratory Study of Strategic Persistence Following Radical Environmental Change. *Academy of Management Journal* 43:837–853.

Augier, M., and M. Prietula. 2007. Perspective—Historical Roots of the A Behavioral Theory of the Firm Model at GSIA. *Organization Science* 18:507–522.

Bailey, M. J. 2006. More Power to the Pill: The Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply. *Quarterly Journal of Economics* 121:289–320.

Baldiga, N. R., and K. B. Coffman. 2018. Laboratory Evidence on the Effects of Sponsorship on the Competitive Preferences of Men and Women. *Management Science* 64:888–901.

Banziger, T., and K. R. Scherer. 2005. The Role of Intonation in Emotional Expressions. *Speech Communication* 46:252–267.

Barabba, V. 2011. *The Decision Loom: A Design for Interactive Decision-Making in Organizations*. Triarchy Press.

Baron, J. N., M. T. Hannan, and M. D. Burton. 1999. Building the Iron Cage: Determinants of Managerial Intensity in the Early Years of Organizations. *American Sociological Review* 64:527–547.

Barr, D. J. 2003. Paralinguistic Correlates of Conceptual Structure. *Psychonomic Bulletin & Review* 10:462–467.

Barr, P. S., J. L. Stimpert, and A. S. Huff. 1992. Cognitive Change, Strategic Action, and Organizational Renewal. *Strategic Management Journal* 13:15–36.

Barry, A. 2008. *The Form, Function And Distribution Of High Rising Intonation*. Riga, Latvia: VDM Verlag Dr. Müller.

Barua, A., L. F. Davidson, D. V. Rama, and S. Thiruvadi. 2010. CFO Gender and Accruals Quality. *Accounting Horizons* 24:25–39.

Baum, J. A., T. J. Rowley, A. V. Shipilov, and Y.-T. Chuang. 2005. Dancing with Strangers: Aspiration Performance and the Search for Underwriting Syndicate Partners. *Administrative Science Quarterly* 50:536–575.

Bechky, B. A. 2003. Sharing Meaning across Occupational Communities: The Transformation of Understanding on a Production Floor. *Organization Science* 14:312–330.

Beckman, C. M. 2006. The Influence of Founding Team Company Affiliations on Firm Behavior. *Academy of Management Journal* 49:741–758.

Begeny, C. T., M. K. Ryan, C. A. Moss-Racusin, and G. Ravetz. 2020. In Some Professions, Women Have Become Well Represented, yet Gender Bias Persists—Perpetuated by Those Who Think It Is Not Happening. *Science Advances* 6:eaba7814.

Berg, N., and D. Lien. 2002. Measuring the Effect of Sexual Orientation on Income: Evidence of Discrimination? *Contemporary Economic Policy* 20:394–414.

Bernard, V. L., and J. K. Thomas. 1990. Evidence That Stock Prices Do Not Fully Reflect the Implications of Current Earnings for Future Earnings. *Journal of Accounting and Economics* 13:305–340.

Bigelow, L., L. Lundmark, J. McLean Parks, and R. Wuebker. 2014. Skirting the Issues: Experimental Evidence of Gender Bias in IPO Prospectus Evaluations. *Journal of Management* 40:1732–1759.

Blackler, F., and S. McDonald. 2000. Power, Mastery and Organizational Learning. *Journal of Management Studies* 37:833–852.

Blau, B. M., J. R. DeLisle, and S. M. Price. 2015. Do Sophisticated Investors Interpret Earnings Conference Call Tone Differently than Investors at Large? Evidence from Short Sales. *Journal of Corporate Finance* 31:203–219.

Blau, F. D., and L. M. Kahn. 2017. The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature* 55:789–865.

Boeker, W. 1989. Strategic Change: The Effects of Founding and History. *Academy of Management journal* 32:489–515.

Bohnet, I., A. van Geen, and M. Bazerman. 2016. When Performance Trumps Gender Bias: Joint vs. Separate Evaluation. *Management Science* 62:1225–1234.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer. 2019. Beliefs about Gender. *American Economic Review* 109:739–773.

Boserup, E. 1967. Geography and a Crowding World. In *Symposium on Population Pressures upon Physical and Social Resources in the Developing Lands*. Ed. Wilbur Zelinsky, Leszek A. Kosinski, and R. Mansell Prothero. Oxford: Oxford University Press.

Boukus, E., and J. V. Rosenberg. 2006. The Information Content of FOMC Minutes. Mimeo, Federal Reserve Bank of New York, New York, NY.

Boyle, E., and Z. Shapira. 2012. The Liability of Leading: Battling Aspiration and Survival Goals in the Jeopardy! Tournament of Champions. *Organization Science* 23:1100–1113.

Brennan, S. E., and M. Williams. 1995. The Feeling of Another's Knowing: Prosody and Filled Pauses as Cues to Listeners about the Metacognitive States of Speakers. *Journal of Memory and Language* 34:383–398.

Brown, N. C., R. M. Crowley, and W. B. Elliott. 2020. What Are You Saying? Using Topic to Detect Financial Misreporting. *Journal of Accounting Research* 58:237–291.

Brown, N. C., W. Hu, T. Shohfi, B. B. Francis, T. Zhang, and D. Xin. 2023. Gender and Earnings Conference Calls. Working Paper, University of Illinois, Urbana-Champaign.

Brown, S. L., and K. M. Eisenhardt. 1998. *Competing on the Edge: Strategy as Structured Chaos*. Cambridge, Massachusetts: Harvard Business Press.

Burke, K. 1934. *Permanence and Change: An Anatomy of Purpose*. California: University of California Press.

Burke, K. 1937. *Attitudes toward History: With a New Afterword*. California: University of California Press.

Burton, M. D., J. B. Sørensen, and C. M. Beckman. 2002. Coming from Good Stock: Career Histories and New Venture Formation. In *Social Structure and Organizations Revisited*. Ed. Michael Lounsbury, and Marc Ventresca, pp. 229–262. Leeds, UK: Emerald Publishing Limited.

Butler, J. 1990. *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge.

Calabretta, G., G. Gemser, and N. M. Wijnberg. 2017. The Interplay between Intuition and Rationality in Strategic Decision Making: A Paradox Perspective. *Organization Studies* 38:365–401.

Carlesen, A. 2018. #MeToo Brought Down 201 Powerful Men. Nearly Half of Their Replacements Are Women. *The New York Times* .

Carroll, G. R., and M. T. Hannan. 1989. Density Dependence in the Evolution of Populations of Newspaper Organizations. *American Sociological Review* 54:524–541.

Carter, D. A., F. D'Souza, B. J. Simkins, and W. G. Simpson. 2010. The Gender and Ethnic Diversity of US Boards and Board Committees and Firm Financial Performance. *Corporate Governance: An International Review* 18:396–414.

Carter, D. A., B. J. Simkins, and W. G. Simpson. 2003. Corporate Governance, Board Diversity, and Firm Value. *Financial Review* 38:33–53.

Cattaneo, M. D., R. K. Crump, M. H. Farrell, and Y. Feng. 2022. On Binscatter. Staff Reports 881, Federal Reserve Bank of New York, New York, NY.

Charness, G., and A. Rustichini. 2011. Gender Differences in Cooperation with Group Membership. *Games and Economic Behavior* 72:77–85.

Chatman, J. A., D. F. Caldwell, C. A. O'Reilly, and B. Doerr. 2014. Parsing Organizational Culture: How the Norm for Adaptability Influences the Relationship between Culture Consensus and Financial Performance in High-technology Firms. *Journal of Organizational Behavior* 35:785–808.

Chen, J. V., V. Nagar, and J. Schoenfeld. 2018. Manager-Analyst Conversations in Earnings Conference Calls. *Review of Accounting Studies* 23:1315–1354.

Chen, S., D. Matsumoto, and S. Rajgopal. 2011. Is Silence Golden? An Empirical Analysis of Firms That Stop Giving Quarterly Earnings Guidance. *Journal of Accounting and Economics* 51:134–150.

Chun, M. M., J. D. Golomb, and N. B. Turk-Browne. 2011. A Taxonomy of External and Internal Attention. *Annual Review of Psychology* 62:73–101.

Clark, H. H. 1996. *Using Language*. Cambridge, England: Cambridge University Press.

Clark, H. H. 1998. Communal Lexicons. In *Context in Language Learning and Language Understanding*. Ed. K. Malmkjaer, and J. Williams, pp. 63–87. Cambridge, UK: Cambridge University Press.

Clark, H. H., and D. Wilkes-Gibbs. 1986. Referring as a Collaborative Process. *Cognition* 22:1–39.

Clopper, C. G., and R. Smiljanic. 2011. Effects of Gender and Regional Dialect on Prosodic Patterns in American English. *Journal of Phonetics* 39:237–245.

Coffman, K., C. B. Flikkema, and O. Shurchkov. 2021. Gender Stereotypes in Deliberation and Team Decisions. *Games and Economic Behavior* 129:329–349.

Coffman, K. B., C. L. Exley, and M. Niederle. 2021. The Role of Beliefs in Driving Gender Discrimination. *Management Science* 67:3551–3569.

Colunga, E., and L. B. Smith. 2005. From the Lexicon to Expectations about Kinds: A Role for Associative Learning. *Psychological Review* 112:347–382.

Colyvas, J. A., and W. W. Powell. 2006. Roads to Institutionalization: The Remaking of Boundaries between Public and Private Science. *Research in Organizational Behavior* 27:305–353.

Conklin, M. 2020. #MeToo Effects on Juror Decision Making. *Calif. L. Rev. Online* 11:179.

Conley, J. M., W. M. O'Barr, and E. A. Lind. 1979. The Power of Language: Presentational Style in the Courtroom. *Duke Law Journal* 1978:1375–1399.

Contu, A., and H. Willmott. 2003. Re-Embedding Situatedness: The Importance of Power Relations in Learning Theory. *Organization Science* 14:283–296.

Cook, A., A. Esplin, C. Glass, J. S. Judd, and K. Olsen. 2019. Management Forecasts, Analyst Revisions, and Investor Reactions: The Effect of CEO Gender. Discussion paper, Utah State University, Utah.

Cooper, A. C., and D. Schendel. 1976. Strategic Responses to Technological Threats. *Business Horizons* 19:61–69.

Coopey, J., and J. Burgoyne. 2000. Politics and Organizational Learning. *Journal of Management Studies* 37:869–886.

Corritore, M., A. Goldberg, and S. B. Srivastava. 2020. Duality in Diversity: How Intrapersonal and Interpersonal Cultural Heterogeneity Relate to Firm Performance. *Administrative Science Quarterly* 65:359–394.

Cowen, A. P., and N. V. Montgomery. 2020. To Be or Not to Be Sorry? How CEO Gender Impacts the Effectiveness of Organizational Apologies. *Journal of Applied Psychology* 105:196.

Cramton, C. D. 2001. The Mutual Knowledge Problem and Its Consequences for Dispersed Collaboration. *Organization Science* 12:346–371.

Cremer, J., L. Garicano, and A. Prat. 2007. Language and the Theory of the Firm. *The Quarterly Journal of Economics* 122:373–407.

Cyert, R. M., and J. G. March. 1963. *A Behavioral Theory of the Firm*. Prentice Hall.

Daft, R. L., and K. E. Weick. 1984. Toward a Model of Organizations as Interpretation Systems. *Academy of Management Review* 9:284–295.

Dalpiaz, E., and G. Di Stefano. 2018. A Universe of Stories: Mobilizing Narrative Practices during Transformative Change. *Strategic Management Journal* 39:664–696.

Daly, N., and P. Warren. 2001. Pitching It Differently in New Zealand English: Speaker Sex and Intonation Patterns. *Journal of Sociolinguistics* 5:85–96.

De Amicis, C., S. Falconieri, and M. Tastan. 2021. Sentiment Analysis and Gender Differences in Earnings Conference Calls. *Journal of Corporate Finance* 71.

de Beauvoir, S. 1949. *The Second Sex*. Vintage Classics.

Dearborn, D. C., and H. A. Simon. 1958. Selective Perception: A Note on the Departmental Identifications of Executives. *Sociometry* 21:140–144.

Deephouse, D. L. 1996. Does Isomorphism Legitimate? *Academy of Management Journal* 39:1024–1039.

Denrell, J., C. Fang, and D. A. Levinthal. 2004. From T-mazes to Labyrinths: Learning from Model-Based Feedback. *Management Science* 50:1366–1378.

DesJardine, M., P. Bansal, and Y. Yang. 2019. Bouncing Back: Building Resilience through Social and Environmental Practices in the Context of the 2008 Global Financial Crisis. *Journal of Management* 45:1434–1460.

Dezso, C. L., and D. G. Ross. 2012. Does Female Representation in Top Management Improve Firm Performance? A Panel Data Investigation. *Strategic Management Journal* 33:1072–1089.

DiMaggio, P., and W. Powell. 1983. The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review* 48:147–160.

Dimov, D., P. M. de Holan, and H. Milanov. 2012. Learning Patterns in Venture Capital Investing in New Industries. *Industrial and Corporate Change* 21:1389–1426.

Divakaruni, A., L. Fritsch, H. Jones, and A. D. Morrison. 2023. Market Reactions to Gendered Speech Patterns: Uptalk, Earnings Calls, and the #MeToo Movement. Working Paper, University of Oxford, Oxford, UK.

Dixon-Fowler, H. R., A. E. Ellstrand, and J. L. Johnson. 2013. Strength in Numbers or Guilt by Association? Intragroup Effects of Female Chief Executive Announcements. *Strategic Management Journal* 34:1488–1501.

Dobbin, F. 1994. Cultural Models of Organization: The Social Construction of Rational Organizing Principles. In *The Sociology of Culture: Emerging Theoretical Perspectives*. Ed. Diana Crane, pp. 117–141. Oxford, UK: Blackwell Publishing Ltd.

Dobbin, F., and T. J. Dowd. 1997. How Policy Shapes Competition: Early Railroad Foundings in Massachusetts. *Administrative Science Quarterly* 42:501–529.

Dobrev, S. D., T.-Y. Kim, and G. R. Carroll. 2003. Shifting Gears, Shifting Niches: Organizational Inertia and Change in the Evolution of the US Automobile Industry, 1885–1981. *Organization Science* 14:264–282.

Dobrev, S. D., T.-Y. Kim, and M. T. Hannan. 2001. Dynamics of Niche Width and Resource Partitioning. *American Journal of Sociology* 106:1299–1337.

Dokko, G., S. L. Wilk, and N. P. Rothbard. 2009. Unpacking Prior Experience: How Career History Affects Job Performance. *Organization Science* 20:51–68.

Dosi, G., and Y. Kaniovski. 1994. On "Badly Behaved" Dynamics. *Journal of Evolutionary Economics* 4:93–123.

Dosi, G., M. C. Pereira, A. Roventini, and M. E. Virgillito. 2017. When More Flexibility Yields More Fragility: The Microfoundations of Keynesian Aggregate Unemployment. *Journal of Economic Dynamics and Control* 81:162–186.

Druz, M., A. F. Wagner, and R. Zeckhauser. 2016. *Reading Managerial Tone: How Analysts and the Market Respond to Conference Calls*. Harvard Kennedy School, John F. Kennedy School of Government.

Dunn, M. B., and C. Jones. 2010. Institutional Logics and Institutional Pluralism: The Contestation of Care and Science Logics in Medical Education, 1967–2005. *Administrative Science Quarterly* 55:114–149.

Eckert, P., and S. McConnell-Ginet. 2003. *Language and Gender*. Cambridge: Cambridge University Press.

Eggers, J. P., and S. Kaplan. 2013. Cognition and Capabilities: A Multi-Level Perspective. *Academy of Management Annals* 7:295–340.

Even-Tov, O., K. Huang, and B. Trueman. 2020. Call Me by Your Name: The Effect of Analyst-CEO First Name Commonality on Analyst Forecast Accuracy. Working Paper, Hass School of Business, University of California, Berkeley.

Faccio, M., M.-T. Marchica, and R. Mura. 2016. CEO Gender, Corporate Risk-Taking, and the Efficiency of Capital Allocation. *Journal of Corporate Finance* 39:193–209.

Fang, C., and D. Levinthal. 2009. Near-Term Liability of Exploitation: Exploration and Exploitation in Multistage Problems. *Organization Science* 20:538–551.

Feldman, M. S., and B. T. Pentland. 2003. Reconceptualizing Organizational Routines as a Source of Flexibility and Change. *Administrative Science Quarterly* 48:91–118.

Fiol, C. M., and M. A. Lyles. 1985. Organizational Learning. *Academy of Management Review* 10:803–813.

Fligstein, N. 1996. Markets as Politics: A Political-cultural Approach to Market Institutions. *American Sociological Review* 61:656–673.

Fox, L. 2003. *Enron: The Rise and Fall*. John Wiley and Sons.

Francis, B. B., I. Hasan, Y. V. Shen, and Q. Wu. 2021. Do Activist Hedge Funds Target Female CEOs? The Role of CEO Gender in Hedge Fund Activism. *Journal of Financial Economics* 141:372–393.

Francoeur, C., R. Labelle, and B. Sinclair-Desgagné. 2008. Gender Diversity in Corporate Governance and Top Management. *Journal of Business Ethics* 81:83–95.

Francoeur, C., Y. Li, Z. Singer, and J. Zhang. 2022. Earnings Forecasts of Female CEOs: Quality and Consequences. *Review of Accounting Studies* pp. 1–44.

Frege, G. 1948. Sense and Reference. *The Philosophical Review* 57:209–230.

Friedan, B. 1963. *The Feminine Mystique*. London: Penguin Modern Classics.

Fussell, S. R., and R. M. Krauss. 1992. Coordination of Knowledge in Communication: Effects of Speakers' Assumptions about What Others Know. *Journal of Personality and Social Psychology* 62:378.

Gallie, W. B. 1956. Art as an Essentially Contested Concept. *The Philosophical Quarterly (1950-)* 6:97–114.

Galor, O., Ö. Özak, and A. Sarid. 2020. Linguistic Traits and Human Capital Formation. *AEA Papers and Proceedings* 110:309–313.

Garud, R., and M. A. Rappa. 1994. A Socio-Cognitive Model of Technology Evolution: The Case of Cochlear Implants. *Organization Science* 5:344–362.

Gavetti, G. 2005. Cognition and Hierarchy: Rethinking the Microfoundations of Capabilities' Development. *Organization Science* 16:599–617.

Gavetti, G., H. R. Greve, D. A. Levinthal, and W. Ocasio. 2012. The Behavioral Theory of the Firm: Assessment and Prospects. *Academy of Management Annals* 6:1–40.

Gavetti, G., R. Henderson, and S. Giorgi. 2005. Kodak and the Digital Revolution. *Harvard Business School Case Study* .

Gavetti, G., and D. Levinthal. 2000. Looking Forward and Looking Backward: Cognitive and Experiential Search. *Administrative Science Quarterly* 45:113.

Gentner, D. 2010. Bootstrapping the Mind: Analogical Processes and Symbol Systems. *Cognitive Science* 34:752–775.

Giuliano, P. 2015. The Role of Women in Society: From Preindustrial to Modern Times. *CESifo Economic Studies* 61:33–52.

Giuliano, P. 2020. Gender and Culture. *Oxford Review of Economic Policy* 36:26.

Goldberg, A., and S. K. Stein. 2018. Beyond Social Contagion: Associative Diffusion and the Emergence of Cultural Variation. *American Sociological Review* 83:897–932.

Gorodnichenko, Y., T. Pham, and O. Talavera. 2023. The Voice of Monetary Policy. *American Economic Review* 113:548–584.

Gow, I. D., D. F. Larcker, and A. A. Zakolyukina. 2021. Non-answers during Conference Calls. *Journal of Accounting Research* 59:1349–1384.

Greve, H. R. 1995. Jumping Ship: The Diffusion of Strategy Abandonment. *Administrative Science Quarterly* pp. 444–473.

Greve, H. R. 1998. Performance, Aspirations, and Risky Organizational Change. *Administrative Science Quarterly* pp. 58–86.

Greve, H. R. 2003. A Behavioral Theory of R&D Expenditures and Innovations: Evidence from Shipbuilding. *Academy of Management Journal* 46:685–702.

Greve, H. R. 2009. Bigger and Safer: The Diffusion of Competitive Advantage. *Strategic Management Journal* 30:1–23.

Guy, G. R., and J. Vonwiller. 1984. The Meaning of an Intonation in Australian English. *Australian Journal of Linguistics* 4:1–17.

Haleblian, J., J.-Y. Kim, and N. Rajagopalan. 2006. The Influence of Acquisition Experience and Performance on Acquisition Behavior: Evidence from the US Commercial Banking Industry. *Academy of Management Journal* 49:357–370.

Hancock, A., L. Colton, and F. Douglas. 2014. Intonation and Gender Perception: Applications for Transgender Speakers. *Journal of Voice* 28:203–209.

Hannan, M. T. 2010. Partiality of Memberships in Categories and Audiences. *Annual Review of Sociology* 36:159–181.

Hansen, M. T., and M. R. Haas. 2001. Competing for Attention in Knowledge Markets: Electronic Document Dissemination in a Management Consulting Company. *Administrative Science Quarterly* 46:1–28.

Hansen, S., M. McMahon, and A. Prat. 2018. Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach. *Quarterly Journal of Economics* 133:801–870.

Hassan, T. A., S. Hollander, L. van Lent, and A. Tahoun. 2019. Firm-Level Political Risk: Measurement and Effects. *Quarterly Journal of Economics* 134:2135–2202.

Hazenberg, E. N. L. 2012. Language and Identity Practice: A Sociolinguistic Study of Gender in Ottawa, Ontario. Ph.D. thesis, Memorial University of Newfoundland.

Hazenberg, E. N. L. 2013. "Is This It?" Stance and Indexicality in Transsexual Speech. In *4th New Zealand Discourse Conference*, Auckland University of Technology.

Heimeriks, K. H., M. Schijven, and S. Gates. 2012. Manifestations of Higher-Order Routines: The Underlying Mechanisms of Deliberate Learning in the Context of Postacquisition Integration. *Academy of Management Journal* 55:703–726.

Helfat, C. E. 1997. Know-how and Asset Complementarity and Dynamic Capability Accumulation: The Case of R&D. *Strategic Management Journal* 18:339–360.

Henderson, R. M., and K. B. Clark. 1990. Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms. *Administrative Science Quarterly* 35:9–30.

Henisz, W. J., and A. Delios. 2001. Uncertainty, Imitation, and Plant Location: Japanese Multinational Corporations, 1990-1996. *Administrative Science Quarterly* 46:443–475.

Henton, C. 1995. Pitch Dynamism in Female and Male Speech. *Language & Communication* 15:43–61.

Henton, C. G. 1989. Fact and Fiction in the Description of Female and Male Pitch. *Language & Communication* 9:299–311.

Higgins, M. C. 2005. *Career Imprints: Creating Leaders across an Industry*. , vol. 16. New Jersey: John Wiley & Sons.

Hoffman, A. J., and W. Ocasio. 2001. Not All Events Are Attended Equally: Toward a Middle-Range Theory of Industry Attention to External Events. *Organization Science* 12:414–434.

Hollander, S., M. Pronk, and E. Roelofsen. 2010. Does Silence Speak? An Empirical Analysis of Disclosure Choices during Conference Calls. *Journal of Accounting Research* 48:531–563.

Hoobler, J. M., C. R. Masterson, S. M. Nkomo, and E. J. Michel. 2018. The Business Case for Women Leaders: Meta-analysis, Research Critique, and Path Forward. *Journal of Management* 44:2473–2499.

Huang, J., and D. J. Kisgen. 2013. Gender and Corporate Finance: Are Male Executives Overconfident Relative to Female Executives? *Journal of Financial Economics* 108:822–839.

Huber, G. P. 1991. Organizational Learning: The Contributing Processes and the Literatures. *Organization Science* 2:88–115.

Iyer, D. N., and K. D. Miller. 2008. Performance Feedback, Slack, and the Timing of Acquisitions. *Academy of Management Journal* 51:808–822.

Jacobides, M. G. 2007. The Inherent Limits of Organizational Structure and the Unfulfilled Role of Hierarchy: Lessons from a near-War. *Organization Science* 18:455–477.

Jensen, R. J., and G. Szulanski. 2007. Template Use and the Effectiveness of Knowledge Transfer. *Management Science* 53:1716–1730.

Jeong, S.-H., and D. A. Harrison. 2017. Glass Breaking, Strategy Making, and Value Creating: Meta-analytic Outcomes of Women as CEOs and TMT Members. *Academy of Management Journal* 60:1219–1252.

Jiang, F., J. Lee, X. Martin, and G. Zhou. 2019. Manager Sentiment and Stock Returns. *Journal of Financial Economics* 132:126–149.

Joenväärä, J., J. Karppinen, M. Teo, and C. I. Tiu. 2019. Text Sophistication and Sophisticated Investors. Working Paper, Aalto University, Aalto.

Johnson, E. N., L. A. Kidwell, D. J. Lowe, and P. M. J. Reckers. 2017. Who Follows the Unethical Leader? The Association Between Followers' Personal Characteristics and Intentions to Comply in Committing Organizational Fraud. *Journal of Business Ethics* 154:181–193.

Johnson, S., K. Keplinger, J. Kirk, and L. Barnes. 2019. Has Sexual Harassment at Work Decreased Since #MeToo? *Harvard Business Review* .

Johnson, V. 2007. What Is Organizational Imprinting? Cultural Entrepreneurship in the Founding of the Paris Opera. *American Journal of Sociology* 113:97–127.

Kale, P., and H. Singh. 2007. Building Firm Capabilities through Learning: The Role of the Alliance Learning Process in Alliance Capability and Firm-level Alliance Success. *Strategic Management Journal* 28:981–1000.

Kaplan, S. 2008. Framing Contests: Strategy Making under Uncertainty. *Organization science* 19:729–752.

Kaplan, S. 2011. Research in Cognition and Strategy: Reflections on Two Decades of Progress and a Look to the Future. *Journal of Management Studies* 48:665–695.

Kaplan, S. N., and L. Zingales. 1997. Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints? *Quarterly Journal of Economics* 112:169–215.

Kauffman, S. 2002. *Investigations*. Oxford, New York: Oxford University Press.

Ke, Z. T., B. T. Kelly, and D. Xiu. 2019. Predicting Returns with Text Data. Working Paper 26186, NBER, Cambridge, MA.

Kelly, D., and T. L. Amburgey. 1991. Organizational Inertia and Momentum: A Dynamic Model of Strategic Change. *Academy of Management Journal* 34:591–612.

Kennedy, M. T. 2008. Getting Counted: Markets, Media, and Reality. *American Sociological Review* 73:270–295.

Ketokivi, M., and X. Castaner. 2004. Strategic Planning as an Integrative Device. *Administrative Science Quarterly* 49:337–365.

Khatri, N., and H. A. Ng. 2000. The Role of Intuition in Strategic Decision Making. *Human Relations* 53:57–86.

Kiesler, S., and L. Sproull. 1982. Managerial Response to Changing Environments: Perspectives on Problem Sensing from Social Cognition. *Administrative Science Quarterly* 27:548–570.

Kim, D., and L. T. Starks. 2016. Gender Diversity on Corporate Boards: Do Women Contribute Unique Skills? *American Economic Review* 106:267–71.

Kimberly, J. R. 1979. Issues in the Creation of Organizations: Initiation, Innovation, and Institutionalization. *Academy of Management Journal* 22:437–457.

Kimberly, J. R., and H. Bouchikhi. 1995. The Dynamics of Organizational Development and Change: How the Past Shapes the Present and Constrains the Future. *Organization Science* 6:9–18.

Klein, N. 1999. *No Logo*. Fourth Estate.

Kocak, Ö., and P. Puranam. 2018. Designing a Culture of Collaboration: When Changing Beliefs Is (Not) Enough. In *Organization Design (Advances in Strategic Management)*. Ed. John Joseph, Oliver Baumann, Richard Burton, and Kannan Srikanth, pp. 27–52. Leeds: Emerald Publishing Limited.

Kocak, Ö., and P. Puranam. 2023. Decoding Culture: Tools for Behavioral Strategists. *Strategy Science* 0.

Kunda, G. 1995. Engineering Culture: Control and Commitment in a High-Tech Corporation. *Organization Science* 6:228–230.

Kunze, A., and A. R. Miller. 2017. Women Helping Women? Evidence from Private Sector Data on Workplace Hierarchies. *Review of Economics and Statistics* 99:769–775.

Lakoff, R., and R. T. Lakoff. 2004. *Language and Woman's Place: Text and Commentaries*. , vol. 3. Oxford University Press, USA.

Lamont, O., C. Polk, and J. Saaá-Requejo. 2001. Financial Constraints and Stock Returns. *Review of Financial Studies* 14:529–554.

Lant, T., and Z. Shapira. 2008. Managerial Reasoning about Aspirations and Expectations. *Journal of Economic Behavior & Organization* 66:60–73.

Lant, T. K. 1992. Aspiration Level Adaptation: An Empirical Exploration. *Management Science* 38:623–644.

Larcker, D. F., and A. A. Zakolyukina. 2012. Detecting Deceptive Discussions in Conference Calls. *Journal of Accounting Research* 50:495–540.

Lee, K., and J. M. Pennings. 2002. Mimicry and the Market: Adoption of a New Organizational Form. *Academy of Management Journal* 45:144–162.

Lee, P. M., and E. H. James. 2007. She'-e-os: Gender Effects and Investor Reactions to the Announcements of Top Executive Appointments. *Strategic Management Journal* 28:227–241.

Levinthal, D., and J. G. March. 1981. A Model of Adaptive Organizational Search. *Journal of Economic Behavior & Organization* 2:307–333.

Levinthal, D., and C. Rerup. 2006. Crossing an Apparent Chasm: Bridging Mindful and Less-Mindful Perspectives on Organizational Learning. *Organization Science* 17:502–513.

Levinthal, D. A. 1997. Adaptation on Rugged Landscapes. *Management Science* 43:934–950.

Levinthal, D. A., and J. G. March. 1993. The Myopia of Learning. *Strategic Management Journal* 14:95–112.

Levinthal, D. A., and C. Rerup. 2021. The Plural of Goal: Learning in a World of Ambiguity. *Organization Science* 32:527–543.

Levitt, B., and J. G. March. 1988. Organizational Learning. *Annual Review of Sociology* 14:319–338.

Levon, E. 2016. Gender, Interaction and Intonational Variation: The Discourse Functions of High Rising Terminals in London. *Journal of Sociolinguistics* 20:133–163.

Levon, E., and Y. Ye. 2020. Language, Indexicality and Gender Ideologies: Contextual Effects on the Perceived Credibility of Women. *Gender and Language* 14:123–151.

Levy, R., and M. Mattsson. 2020. The Effects of Social Movements: Evidence From# MeToo. Working Paper, Tel Aviv University, Tel Aviv.

Li, A., S.-c. S. Chiu, D. T. Kong, R. Cropanzano, and C.-W. Ho. 2021. How CEOs Respond to Mortality Salience during the COVID-19 Pandemic: Integrating Terror Management Theory with Regulatory Focus Theory. *Journal of Applied Psychology* 106:1188–1201.

Li, X., R. N. Sullivan, D. Xu, and G. Gao. 2013. Sell-Side Analysts and Gender: A Comparison of Performance, Behavior, and Career Outcomes. *Financial Analysts Journal* 69:83–94.

Lim, E. N., and B. T. Mccann. 2013. The Influence of Relative Values of Outside Director Stock Options on Firm Strategic Risk from a Multiagent Perspective. *Strategic Management Journal* 34:1568–1590.

Lins, K., L. Roth, H. Servaes, and A. M. Tamayo. 2019. Does Corporate Culture Add Value? Evidence from the Harvey Weinstein Scandal and the #MeToo Movement. Working Paper, University of Utah, Utah.

Lins, K. V., L. Roth, H. Servaes, and A. Tamayo. 2020. Gender, Culture, and Firm Value: Evidence from the Harvey Weinstein Scandal and The# MeToo Movement. Discussion Paper DP14703, CEPR, London.

Locke, J. 1847. *An Essay Concerning Human Understanding*. Kay & Troutman.

Loewenstein, J., and D. Gentner. 2005. Relational Language and the Development of Relational Mapping. *Cognitive Psychology* 50:315–353.

Loewenstein, J., W. Ocasio, and C. Jones. 2012. Vocabularies and Vocabulary Structure: A New Approach Linking Categories, Practices, and Institutions. *The Academy of Management Annals* 6:41–86.

Loh, W.-Y. 2011. Classification and Regression Trees. *Wiley interdisciplinary reviews: Data mining and knowledge discovery* 1:14–23.

López-Moctezuma, G. 2016. Sequential Deliberation in Collective Decision-Making: The Case of the FOMC. Discussion Paper Working Paper, Caltech, California.

Loughran, T., and B. McDonald. 2014. Measuring Readability in Financial Disclosures. *Journal of Finance* 69:1643–1671.

Loughran, T., B. McDonald, and I. Pragidis. 2019. Assimilation of Oil News into Prices. *International Review of Financial Analysis* 63:105–118.

Loviglio, J. 2008. Sound Effects: Gender, Voice and the Cultural Work of NPR. *Radio Journal* 5:67–81.

Lowry, O. 2011. Belfast Intonation and Speaker Gender. *Journal of English Linguistics* 39:209–232.

Loyd, D., K. W. Phillips, J. Whitson, and M. C. Thomas-Hunt. 2010. Expertise in Your Midst: How Congruence between Status and Speech Style Affects Reactions to Unique Knowledge. *Group Processes & Intergroup Relations* 13:379–395.

Lucas, H. C., and J. M. Goh. 2009. Disruptive Technology: How Kodak Missed the Digital Photography Revolution. *The Journal of Strategic Information Systems* 18:46–55.

Luo, H., and L. Zhang. 2021. Measuring the Impact of #MeToo on Gender Equity in Hollywood. *Harvard Business Review* .

MacPherson, W. 1999. *The Stephen Lawrence Inquiry*. London: House of Commons.

Mah, J., K. D. Kolev, G. McNamara, L. Pan, and C. E. Devers. 2023. Women in the C-Suite: A Review of Predictors, Experiences, and Outcomes. *Academy of Management Annals* 17:586–625.

Malcolm, N. 2001. *Ludwig Wittgenstein: A Memoir*. Oxford, New York: Oxford University Press.

Malt, B. C., S. A. Sloman, and S. P. Gennari. 2003. Universality and Language Specificity in Object Naming. *Journal of Memory and Language* 49:20–42.

Manela, A., and A. Moreira. 2017. News Implied Volatility and Disaster Concerns. *Journal of Financial Economics* 123:137–162.

Manning, A., and F. Saidi. 2010. Understanding the Gender Pay Gap: What's Competition Got to Do with It? *Cornell ILR Review* 63:681–698.

March, J., and J. Olsen. 1976. *Ambiguity and Choice in Organizations*. Aschehoug AS.

March, J. G. 1962. The Business Firm as a Political Coalition. *The Journal of Politics* 24:662–678.

March, J. G. 1981. Footnotes to Organizational Change. *Administrative Science Quarterly* 26:563–577.

March, J. G. 1988. Variable Risk Preferences and Adaptive Aspirations. *Journal of Economic Behavior & Organization* 9:5–24.

March, J. G. 1991. Exploration and Exploitation in Organizational Learning. *Organization science* 2:71–87.

March, J. G. 1994. *Primer on Decision Making: How Decisions Happen*. New York, NY: Free Press.

March, J. G., M. Schulz, and X. Zhou. 2000. *The Dynamics of Rules: Change in Written Organizational Codes*. Stanford, CA: Stanford University Press.

March, J. G., and H. A. Simon. 1958. *Organizations*. New York: John Wiley & Sons.

Marchetti, A., and P. Puranam. 2022. Organizational Cultural Strength as the Negative Cross-Entropy of Mindshare: A Measure Based on Descriptive Text. *Humanities and Social Sciences Communications* 9:1–14.

Marianne, B. 2011. New Perspectives on Gender. In *Handbook of Labor Economics*. Ed. David Card, and Orley Ashenfelter, pp. 1543–1590. Elsevier.

Marquis, C. 2003. The Pressure of the Past: Network Imprinting in Intercorporate Communities. *Administrative Science Quarterly* 48:655–689.

Marquis, C., and Z. Huang. 2010. Acquisitions as Exaptation: The Legacy of Founding Institutions in the US Commercial Banking Industry. *Academy of Management Journal* 53:1441–1473.

Marquis, C., and A. Tilcsik. 2013. Imprinting: Toward a Multilevel Theory. *Academy of Management Annals* 7:195–245.

Martins, L. L., V. P. Rindova, and B. E. Greenbaum. 2015. Unlocking the Hidden Value of Concepts: A Cognitive Approach to Business Model Innovation. *Strategic Entrepreneurship Journal* 9:99–117.

Matsa, D. A., and A. R. Miller. 2011. Chipping Away at the Glass Ceiling: Gender Spillovers in Corporate Leadership. *American Economic Review* 101:635–39.

Maula, M., K. H. Heimeriks, and T. Keil. 2023. Organizational Experience and Performance: A Systematic Review and Contingency Framework. *Academy of Management Annals* 17:546–585.

Mayer, K. J., and N. S. Argyres. 2004. Learning to Contract: Evidence from the Personal Computer Industry. *Organization Science* 15:394–410.

Mayew, W. J., and M. Venkatachalam. 2012. The Power of Voice: Managerial Affective States and Future Firm Performance. *The Journal of Finance* 67:1–43.

McConnell-Ginet, S. 1978. Intonation in a Man's World. *Signs: Journal of Women in Culture and Society* 3:541–559.

McEvily, B., J. Jaffee, and M. Tortoriello. 2012. Not All Bridging Ties Are Equal: Network Imprinting and Firm Growth in the Nashville Legal Industry, 1933–1978. *Organization Science* 23:547–563.

Melone, N. P. 1994. Reasoning in the Executive Suite: The Influence of Role/Experience-Based Expertise on Decision Processes of Corporate Executives. *Organization Science* 5:438–455.

Merleau-Ponty, M. 1945. *Phénoménologie de La Perception [Phenomenology of Perception]*. Paris: Gallimard.

Mill, J. S. 1869. *The Subjection of Women*. Oxford, United Kingdom: Routledge.

Miller, D., and M.-J. Chen. 1994. Sources and Consequences of Competitive Inertia: A Study of the US Airline Industry. *Administrative Science Quarterly* 39:1–23.

Miller, K. D., and W.-R. Chen. 2004. Variable Organizational Risk Preferences: Tests of the March-Shapira Model. *Academy of Management Journal* 47:105–115.

Mills, C. W. 1939. Language, Logic, and Culture. *American Sociological Review* 4:670–680.

Mills, C. W. 1940. Situated Actions and Vocabularies of Motive. *American Sociological Review* 5:904–913.

Mohr, J. W., and F. Guerra-Pearson. 2010. The Duality of Niche and Form: The Differentiation of Institutional Space in New York City, 1888–1917. In *Categories in Markets: Origins and Evolution*. Emerald Group Publishing Limited.

Mueller, H., and C. Rauh. 2018. Reading Between the Lines: Prediction of Political Violence Using Newspaper Text. *American Political Science Review* 112:358–375.

Nalikka, A. 2009. Impact of Gender Diversity on Voluntary Disclosure in Annual Reports. *Accounting & Taxation* 1:101–113.

Narduzzo, A., E. Rocco, and M. Warglien. 2001. Talking About Routines in the Field: The Emergence of Organizational Capabilities in a New Cellular Phone Network. In *The Nature and Dynamics of Organizational Capabilities*, p. 27. Oxford, New York: Oxford University Press.

Nekhili, M., H. Chakroun, and T. Chtioui. 2018. Women's Leadership and Firm Performance: Family versus Nonfamily Firms. *Journal of Business Ethics* 153:291–316.

Nelson, R. R., and S. G. Winter. 1982. The Schumpeterian Tradeoff Revisited. *The American Economic Review* 72:114–132.

Nigam, A., and W. Ocasio. 2010. Event Attention, Environmental Sensemaking, and Change in Institutional Logics: An Inductive Analysis of the Effects of Public Attention to Clinton's Health Care Reform Initiative. *Organization Science* 21:823–841.

Ocasio, W. 1997. Towards an Attention-based View of the Firm. *Strategic Management Journal* 18:187–206.

Ocasio, W. 2011. Attention to Attention. *Organization Science* 22:1286–1296.

Ocasio, W., and J. Joseph. 2005. Cultural Adaptation and Institutional Change: The Evolution of Vocabularies of Corporate Governance, 1972–2003. *Poetics* 33:163–178.

Ocasio, W., and J. Joseph. 2008. Rise and Fall-or Transformation?: The Evolution of Strategic Planning at the General Electric Company, 1940–2006. *Long Range Planning* 41:248–272.

Ocasio, W., T. Laamanen, and E. Vaara. 2018. Communication and Attention Dynamics: An Attention-Based View of Strategic Change. *Strategic Management Journal* 39:155–167.

Ocasio, W., J. Loewenstein, and A. Nigam. 2015. How Streams of Communication Reproduce and Change Institutional Logics: The Role of Categories. *Academy of Management Review* 40:28–48.

Ocasio, W., L. Rhee, and D. Milner. 2020. Attention, Knowledge, and Organizational Learning. In *The Oxford Handbook of Group and Organizational Learning*. Ed. Linda Argote, and John M. Levine, pp. 81–94. Oxford, New York: Oxford University Press.

Okhuysen, G. A., and B. A. Bechky. 2009. 10 Coordination in Organizations: An Integrative Perspective. *Academy of Management annals* 3:463–502.

Palvia, A., E. Vähämaa, and S. Vähämaa. 2015. Are Female CEOs and Chairwomen More Conservative and Risk Averse? Evidence from the Banking Industry during the Financial Crisis. *Journal of Business Ethics* 131:577–594.

Papadakis, V. M., S. Lioukas, and D. Chambers. 1998. Strategic Decision-Making Processes: The Role of Management and Context. *Strategic Management Journal* 19:115–147.

Parton, S. R., S. A. Siltanen, L. A. Hosman, and J. Langenderfer. 2002. Employment Interview Outcomes and Speech Style Effects. *Journal of Language and Social Psychology* 21:144–161.

Peni, E., and S. Vähämaa. 2010. Female Executives and Earnings Management. *Managerial Finance* 36:629–645.

Phillips, D. J. 2005. Organizational Genealogies and the Persistence of Gender Inequality: The Case of Silicon Valley Law Firms. *Administrative Science Quarterly* 50:440–472.

Pletzer, J. L., R. Nikolova, K. K. Kedzior, and S. C. Voelpel. 2015. Does Gender Matter? Female Representation on Corporate Boards and Firm Financial Performance-a Meta-Analysis. *PloS ONE* 10.

Porac, J. F., H. Thomas, F. Wilson, D. Paton, and A. Kanfer. 1995. Rivalry and the Industry Model of Scottish Knitwear Producers. *Administrative Science Quarterly* 40:203.

Posen, H. E., and J. S. Chen. 2013. An Advantage of Newness: Vicarious Learning despite Limited Absorptive Capacity. *Organization Science* 24:1701–1716.

Posen, H. E., T. Keil, S. Kim, and F. D. Meissner. 2018. Renewing Research on Problemistic Search—A Review and Research Agenda. *Academy of Management Annals* 12:208–251.

Powers, W., R. Troubh, and H. Winokur, Jr. 2002. Report of Investigation by the Special Investigative Committee of the Board of Directors of Enron Corp. Discussion paper, Special Investigation Committee.

Prahalad, C. K., and R. A. Bettis. 1986. The Dominant Logic: A New Linkage between Diversity and Performance. *Strategic Management Journal* 7:485–501.

Raffaelli, R., and C. Snively. 2020. The Reinvention of Kodak. *Harvard Business School Case 419-012* .

Rahmandad, H. 2008. Effect of Delays on Complexity of Organizational Learning. *Management Science* 54:1297–1312.

Reuben, E., P. Sapienza, and L. Zingales. 2014. How Stereotypes Impair Women's Careers in Science. *Proceedings of the National Academy of Sciences* 111:4403–4408.

Rindova, V., E. Dalpiaz, and D. Ravasi. 2011. A Cultural Quest: A Study of Organizational Use of New Cultural Resources in Strategy Formation. *Organization Science* 22:413–431.

Ritchart, A., and A. Arvaniti. 2014. The Form and Use of Uptalk in Southern Californian English. In *Proceedings of the 7th International Conference on Speech Prosody*. Ed. N Campbell, D Gibbon, and D Hirst, pp. 331–335. Dublin, Ireland: International Speech Communication Association.

Rivkin, J. W. 2000. Imitation of Complex Strategies. *Management Science* 46:824–844.

Rivkin, J. W., and N. Siggelkow. 2003. Balancing Search and Stability: Interdependencies Among Elements of Organizational Design. *Management Science* 49:290–311.

Romanelli, E. 1989. Environments and Strategies of Organization Start-up: Effects on Early Survival. *Administrative Science Quarterly* 34:369–387.

Russell, B. 1905. On Denoting. *Mind* 14:479–493.

Santacreu-Vasut, E., A. Shoham, and V. Gay. 2013. Do Female/Male Distinctions in Language Matter? Evidence from Gender Political Quotas. *Applied Economics Letters* 20:495–498.

Scaltsas, T. 1980. The Ship of Theseus. *Analysis* 40:152–157.

Searle, J. 1995. *The Social Construction of Reality*. London: Allen Lane.

Searle, J. 2010. *Making the Social World: The Structure of Human Civilization*. New York: Oxford University Press.

Selznick, P. 1957. *Leadership in Administration: A Sociological Interpretation*. New York, NY: Harper Row.

Senate Subcomittee. 2002. The Role of the Board of Directors in Enron'sCollapse. Report Prepared by the Permanent Subcommittee on Investigations of the Committee on Governmental Affairs. Discussion Paper 107-70, 107th Congress, 2d Session, Washington DC: US Senate.

Seré, M. 2023. Don't Stop Me Now: Gender Attitudes in Academic Seminars Through Machine Learning. Working Paper 23/09, Herman Deleeck Centre for Social Policy, University of Antwerp.

Shinkle, G. A. 2012. Organizational Aspirations, Reference Points, and Goals: Building on the Past and Aiming for the Future. *Journal of Management* 38:415–455.

Siciliano, J. I. 1996. The Relationship of Board Member Diversity to Organizational Performance. *Journal of Business Ethics* 15:1313–1320.

Simon, H. A. 1947. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*. Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization, 3rd Ed. New York, NY, US: Free Press, 3rd edn.

Simon, H. A. 1955. A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics* 69:99–118.

Simon, H. A. 1957. *Models of Man; Social and Rational.* New York: Wiley.

Simon, H. A. 1962. New Developments in the Theory of the Firm. *The American Economic Review* 52:1–15.

Simon, H. A. 1964. On the Concept of Organizational Goal. *Administrative Science Quarterly* 9:1–22.

Simons, T., and P. W. Roberts. 2008. Local and Non-Local Pre-Founding Experience and New Organizational Form Penetration: The Case of the Israeli Wine Industry. *Administrative Science Quarterly* 53:235–265.

Sinha, P. N., P. Jaskiewicz, J. Gibb, and J. G. Combs. 2020. Managing History: How New Zealand's Gallagher Group Used Rhetorical Narratives to Reprioritize and Modify Imprinted Strategic Guideposts. *Strategic Management Journal* 41:557–589.

Smith, E. B., J. Chown, and K. Gaughan. 2021. Better in the Shadows? Public Attention, Media Coverage, and Market Reactions to Female CEO Announcements. *Sociological Science* 8:119–149.

Spike, M., K. Stadler, S. Kirby, and K. Smith. 2017. Minimal Requirements for the Emergence of Learned Signaling. *Cognitive Science* 41:623–658.

Spindler, J. M. 2003. Was That a Question? A Study of High-Rise Terminals. M.A. Thesis, South Dakota State University, South Dakota.

Srikanth, K., and P. Puranam. 2014. The Firm as a Coordination System: Evidence from Software Services Offshoring. *Organization Science* 25:1253–1271.

Srivastava, S. B., A. Goldberg, V. G. Manian, and C. Potts. 2018. Enculturation Trajectories: Language, Cultural Adaptation, and Individual Outcomes in Organizations. *Management Science* 64:1348–1364.

Stinchcombe, A. L. 1965. Social Structure and Organizations. In *Handbook of Organizations*. Ed. J. G. March, pp. 142–193. New York, NY: Routledge.

Suchman, M. C. 1995. Managing Legitimacy: Strategic and Institutional Approaches. *Academy of Management Review* 20:571–610.

Swaminathan, B. 1996. Time-Varying Expected Small Firm Returns and Closed-End Fund Discounts. *The Review of Financial Studies* 9:845–887.

Swidler, A. 1986. Culture in Action: Symbols and Strategies. *American Sociological Review* 51:273.

Talavera, O., S. Yin, and M. Zhang. 2022. Communicating With Voice Emotion-Reciprocal Behavior in Manager-Analyst Interaction. Discussion Paper 22-11, University of Birmingham, Department of Economics.

Tate, G., and L. Yang. 2015. Female Leadership and Gender Equity: Evidence from Plant Closure. *Journal of Financial Economics* 117:77–97.

Teece, D. J. 1986. Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy. *Research Policy* 15:285–305.

Teece, D. J. 1988. Capturing Value from Technological Innovation: Integration, Strategic Partnering, and Licensing Decisions. *Interfaces* 18:46–61.

Tetlock, P. C., M. Saar-Tsechansky, and S. Macskassy. 2008. More Than Words: Quantifying Language to Measure Firms' Fundamentals. *The Journal of Finance* 63:1437–1467.

Thornton, P. H., and W. Ocasio. 1999. Institutional Logics and the Historical Contingency of Power in Organizations: Executive Succession in the Higher Education Publishing Industry, 1958–1990. *American Journal of Sociology* 105:801–843.

Thurstone, L. L. 1919. The Learning Curve Equation. *Psychological Monographs* 26:i–51.

Tilcsik, A. 2014. Imprint–Environment Fit and Performance: How Organizational Munificence at the Time of Hire Affects Subsequent Job Performance. *Administrative Science Quarterly* 59:639–668.

Tomasello, M. 2009. *The Cultural Origins of Human Cognition*. Harvard University Press.

Tomlinson Jr, J. M., and J. E. F. Tree. 2011. Listeners' Comprehension of Uptalk in Spontaneous Speech. *Cognition* 119:58–69.

Tripsas, M., and G. Gavetti. 2017. Capabilities, Cognition, and Inertia: Evidence from Digital Imaging. In *The SMS Blackwell Handbook of Organizational Capabilities*. Ed. Constance E. Helfat, pp. 393–412. New York, NY: Wiley-Blackwell.

Tur-Prats, A. 2019. Family Types and Intimate Partner Violence: A Historical Perspective. *Review of Economics and Statistics* 101:878–891.

Tushman, M. L., and P. Anderson. 2018. Technological Discontinuities and Organizational Environments. In *Organizational Innovation*, pp. 345–372. Oxford, UK: Routledge.

Tushman, M. L., and C. A. O'Reilly III. 1996. Ambidextrous Organizations: Managing Evolutionary and Revolutionary Change. *California Management Review* 38:8–29.

Tyler, B. B., and T. Caner. 2016. New Product Introductions below Aspirations, Slack and R&D Alliances: A Behavioral Perspective. *Strategic Management Journal* 37:896–910.

Vicinanza, P., A. Goldberg, and S. B. Srivastava. 2023. A Deep-Learning Model of Prescient Ideas Demonstrates That They Emerge from the Periphery. *Proceedings of the National Academy of Sciences* 2:1–11.

Wang, G., R. M. Holmes Jr, R. A. Devine, and J. Bishoff. 2018. CEO Gender Differences in Careers and the Moderating Role of Country Culture: A Meta-Analytic Investigation. *Organizational Behavior and Human Decision Processes* 148:30–53.

Warren, P. 2016. *Uptalk: The Phenomenon of Rising Intonation*. Cambridge University Press.

Washburn, M., and P. Bromiley. 2012. Comparing Aspiration Models: The Role of Selective Attention. *Journal of Management Studies* 49:896–917.

Waxman, S. R., and D. B. Markow. 1995. Words as Invitations to Form Categories: Evidence from 12-to 13-Month-Old Infants. *Cognitive Psychology* 29:257–302.

Weber, R. A., and C. F. Camerer. 2003. Cultural Conflict and Merger Failure: An Experimental Approach. *Management Science* 49:400–415.

Wickens, D. D. 1970. Encoding Categories of Words: An Empirical Approach to Meaning. *Psychological Review* 77:1–15.

Wiley, M. G., and A. Eskilson. 1985. Speech Style, Gender Stereotypes, and Corporate Success: What If Women Talk More like Men? *Sex Roles* 12:993–1007.

Wittgenstein, L. 1921. *Tractatus Logico-Philosophicus*. Routledge Classics.

Wittgenstein, L. 1953. *Philosophical Investigations*. New York: John Wiley & Sons.

Wolf, N. 1991. *The Beauty Myth: How Images of Beauty Are Used Against Women*. London: Vintage Classics.

Wollstonecraft, M. 1792. *A Vindication of the Rights of Woman*. Yale University Press.

Yanadori, Y., C. T. Kulik, and J. A. Gould. 2021. Who Pays the Penalty? Implications of Gender Pay Disparities within Top Management Teams for Firm Performance. *Human Resource Management* 60:681–699.

Zander, U., and B. Kogut. 1995. Knowledge and the Speed of the Transfer and Imitation of Organizational Capabilities: An Empirical Test. *Organization Science* 6:76–92.

Zhang, Y., and H. Qu. 2016. The Impact of CEO Succession with Gender Change on Firm Performance and Successor Early Departure: Evidence from China's Publicly Listed Companies in 1997–2010. *Academy of Management Journal* 59:1845–1868.

Zhong, C.-B., J. Loewenstein, and J. K. Murnighan. 2007. Speaking the Same Language: The Cooperative Effects of Labeling in the Prisoner's Dilemma. *Journal of Conflict Resolution* 51:431–456.

Zhu, H., Q. Zhu, and Z. Ding. 2020. The Roles of Chinese CEOs in Managing Individualistic Cultures in Cross-border Mergers and Acquisitions. *Journal of Management Studies* 57:664–697.

Zollo, M. 2009. Superstitious Learning with Rare Strategic Decisions: Theory and Evidence from Corporate Acquisitions. *Organization Science* 20:894–908.

Zollo, M., and H. Singh. 2004. Deliberate Learning in Corporate Acquisitions: Post-acquisition Strategies and Integration Capability in US Bank Mergers. *Strategic Management Journal* 25:1233–1256.

Zollo, M., and S. G. Winter. 2002. Deliberate Learning and the Evolution of Dynamic Capabilities. *Organization Science* 13:339–351.

Zou, X., K.-P. Tam, M. W. Morris, S.-l. Lee, I. Y.-M. Lau, and C.-y. Chiu. 2009. Culture as Common Sense: Perceived Consensus versus Personal Beliefs as Mechanisms of Cultural Influence. *Journal of Personality and Social Psychology* 97:579.